VISIONS OF 4G

(Submitted to IEE Electronics & Communications Engineering Journal, for publication autumn/winter 2000)

B G Evans

Director CCSR, University of Surrey, Guildford GU2 7XH

(Chair of the Mobile VCE Visions Group)

and

K Baughan

VP New Business Development, Nokia Mobile Phones, Summit Avenue, Farnborough, Hants. GU14 0NG

(Chairman of the Mobile VCE Board)

1 Introduction

Pick up any newspaper today and it is a safe bet that you will find an article somewhere relating to Mobile Communications. If it is not in the technology section it will almost certainly be in the business section and relate to the increasing share prices of operators or equipment manufacturers, or acquisitions and take over thereof. Such is the pervasiveness of mobile communications that it is affecting virtually everyone's life; has become a major political topic and a significant contributor to national GDP.

The major driver to change in the mobile area in the last ten years has been the massive enabling implications of digital technology, both in digital signal processing and in service provision. The equivalent driver now, and in the next five years, will be the all pervasiveness of software in both networks and terminals. The digital revolution is well underway and we stand at the doorway to the software revolution. Accompanying these changes are the societal developments involving the extensions in the use of mobiles. Starting out with speech dominated services we are now experiencing massive growth in applications involving SMS (Short Messaging Services) together with the start of internet applications using WAP (Wireless Application Protocol) and i-mode. The mobile phone has not only followed the watch, calculator and organiser as an essential personal attribute - but it has subsumed all of them. With the new internet extensions it will also converge the PC,HiFi and television and provide mobility to erstwhile facilities only available on one network, so called Virtual Home Environment (VHE).

The development from first generation analogue systems (1985) to second generation digital GSM (2G) (1992) was the heart of the digital revolution. But much more than this it was a huge success for standardisation emanating from Europe and gradually spreading globally. (GSM becomes Global System for Mobile Communications).

However, roaming world-wide still presents some problems with pockets of US standard IS-95 (a cdma rather than tdma digital system) and IS-136 (a tdma variant), still endemic in some countries. GSM (2G) extensions via GPRS and EDGE (E-GPRS)as well as WAP and i-mode (so called 2.5G) are the migrations to allow higher data rates as well as speech, prior to the introduction of 3G.

Mobile systems comprise a radio access together with a supporting core network. The latter in GSM is characterised by MAP (Mobile Applications Protocol) which provides the mobility management features of the system.

GSM was designed around digital speech services or for low bit rate data that could fit into a speech channel (eg 9.6 kb/s). It is a circuit rather than a packet oriented network which is inefficient for data communications. In order to address the rapid popularity increase of internet services, GPRS (General Packet Radio Services) is being added to GSM in 2000 to allow packet (IP) communications at up to about 100 kb/s.

Third generation (3G) were standardised in 1999 (IMT-2000 within ITU-R which includes the UMTS European ETSI standard plus US derived CDMA 2000 and

Japanese NTT DoCoMo WCDMA). Such systems extend services to high quality multimedia (multirate) and to convergent networks of fixed, cellular and satellite components. The radio air interface standards are based upon WCDMA (UTRA FDD, UTRA TDD in UMTS and multicarrier CDMA 2000, single carrier UWC-136 on derived US standards). The core network has not been standardised, but a group of three; evolved GSM (MAP), evolved ANSI-41 and IP-based are all candidates. 3G is also about a diversity of terminal types including many non-voice terminals such as those embedded in all sorts of consumer products. Bluetooth (another standard not within the 3G orbit, but likely to be associated with it) is a short range system which addresses such applications. Thus services from a few b/s up to 2Mb/s can be envisioned.

For broadband indoor wireless communications, standards such as HIPERLAN 2 (ETSI/BRAN) and IEEE 802.11a have emerged to support IP-based services and provide some QoS support. Such systems are based on OFDM rather than CDMA and are planned to operate in the 5GHz band.

Whereas 2G operates in 900 and 1800/1900 MHz frequency bands, 3G is intended to operate in wider bandwidth allocations at 2GHz. These new frequency bands will provide wider bandwidths for some multimedia services and the first allocations have occurred in some countries via spectrum auctions (eg. UK, Holland and Germany) or beauty contests (France, Italy). The opportunity has also been taken to increase competition by allowing new operators into the bands as well as extending existing operator licences. These new systems will comprise micro cells as well as macrocells in order to deliver the higher capacity services efficiently. 3G and 2G will continue to co-exist for some time with optimisation of service provision between them. Various modes of delivery will be used to improve coverage in urban, suburban and rural areas with satellite (and possibly HAPS - High Altitude Platform Stations) playing a role. The story of the evolution of mobile radio generations is summed up in Figure 1.

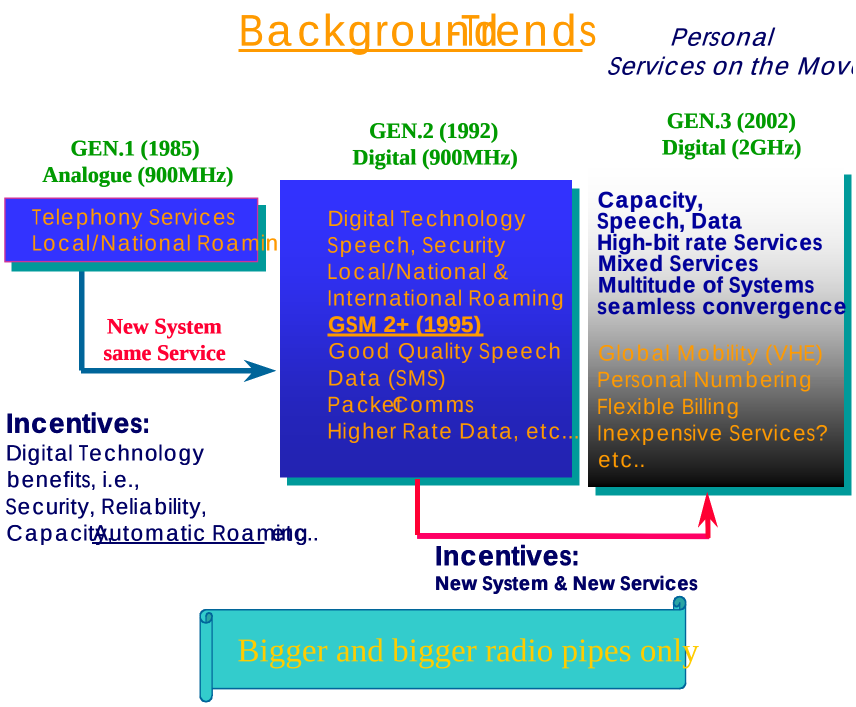

Fig 1: Evolution of mobile radio generations

Already as we move from 2G to 3G the convergence of communications and computing is central to the realisation of the new generation of service and applications. Digital technology enables dynamic adaptation of systems, and intercommunicating software embedded in networks and terminals allows efficient control of the new networks. This is accentuated as we move from 3G to 4G to extend the range and bit rate of services and to consider convergence of fixed, mobile and broadcast networks, service provision and terminal types.

In this paper we introduce the basic ideas and thinking behind the UK Mobile VCEs second phase research programme (1999 - 2003) in the form of "visions for 4G". A Visions Group has been set up to produce and maintain an evolving picture of 4G and to flow down these ideas to the work areas and researchers. The aim is to provide an umbrella vision to harmonise the research work in the various areas.

4G VISIONS GROUP

Barry Evans (University of Surrey)
Walter Tuttlebee (Director, MVCE)
Keith Baughan (Nokia Mobile Phones_
Mike Walker (Vodafone)
Peter Ramsale (On2One)
Dave Thomas (Texas Instruments)

In the next section we explain the limitations of 3G and derive the drivers for 4G. In subsequent sections we then present "the 4G Vision" and some of the research challenges that this presents. The approach that is taken here is one of developing a technical vision. However, it is based upon likely user scenarios which have been developed within the Mobile VCE. (1)

2 Limitations of 3G and Drivers for 4G

From its basic conception to the time of roll-out took around ten years for 2G and a similar period applied to 3G which will commence service in 2001/2 and reach full deployment by 2005. Thus by 2010 it will be time to deploy 4G networks and working backwards with the ten year cycle, it is clear that the year 2000 is appropriate to start with visions for 4G and a research programme aimed at the key issues. The Mobile VCE’s second phase research programme has been constructed to meet this aim.

The starting point was to look at current trends. Here we see a phenomenal growth in mobiles with an estimated global user base that will exceed one billion by 2003. Already mobile communications have exceeded fixed communications in several countries and it is foreseen that mobile will subsume fixed by 2010 (Fixed-mobile convergence will be complete). Currently short messaging is booming, especially with the younger generation, with averages of upwards of 100 messages per month dominating monthly bills. Business take up of SMS via information services is also increasing and providing a start for mobile e-commerce, but is currently very much limited by the available bit rates. This will be improved with the introduction of GPRS. At the time of writing we are at the beginning of the next stage on from short messaging which is to provide an efficient and convenient user interface to enable internet content to be accessed on mobile devices - the "internet to the pocket era". The European introduction of WAP (using WML) has been slow to gain market ground but by contrast in Japan, NTT DoCoMo's "i-mode" system had over 7 million submissions (May 2000) and is picking up 50,000 new customers per day. Customers are already browsing the internet, exchanging email, conducting banking and stock transactions, making flight reservations and checking news and weather via HTML-based text information on their phones. Java is expected to be available on i-mode phones soon, allowing the download of agents, games etc and the introduction of location based services. In Japan, the number of net phones has now passed the number of wired internet customers and is setting the trend that others will surely follow when 3G opens up more bandwidth and improved quality.

Thus 3G will provide a significant step in the evolution of mobile personal communications. Mobility appears to be one of the fundamental elements in the evolution of the information society. As service provision based on "Network Centric architectures is gradually giving way to the "Edge-Centric" architectures access is needed from more and more places at all times. But can 3G deliver?
It is true that 3G can support multimedia internet type services at improved speeds and quality compared to 2G. The WCDMA based air-interface has been designed to provide improved high capacity coverage for medium bit rates (384 kb/s) with limited coverage and with up to 2Mb/s (in indoor environments). The statistical multiplexing on the air also improves the efficiency of packet mode transmission. However, there are limitations with 3G as follows:

- Extension to higher data rates is difficult with CDMA - due to excessive interference between services.
- It is difficult to provide a full range of multirate services, all with different QoS (Quality of Service) and performance requirement - due to the constraints imposed on the core network by the air interference standard, (eg it is not a fully integrated system).

In addition the bandwidth available in the 2GHz bands allocated for 3G will soon become saturated and there are constraints on the combination of frequency and time division duplex modes imposed by regulators, to serve different environments efficiently.

By the year 2010, one of the key enabling technology developments will be embedded radio - the widespread availability and use of the $1 radio chop, which will evolve from short range wireless developments such as Bluetooth. Embedded radio will eventually become as common as embedded microprocessors are today, with perhaps 50 such devices in the typical home, with the user mostly unaware of their presence. As they interact, in response to the user arriving home for example, they will form a Home Area Network (HAN). Similarly such devices will be present in large numbers in vehicles (the Vehicular Area Network, VAN), in personal belongings (the Personal Area Network, PAN), in the public environment, etc. Such chips will serve as a means of short range communication between objects and devices, offering capabilities for monitoring and control, in most cases without user knowledge or intervention.

As a person moves between these environments such short range links will allow his personal profiles and preferences to move with him, with the hotel room automatically configuring itself to his personal preferred temperatures, TV channels/interests, lighting etc. However, the integration of such links with wide area mobile access will enable far more powerful service concepts, as mobile agents access this pervasive network of sensors and information access on the user's behalf to perform and even pre-empt his needs and wishes.

In the 1G to 2G transition, as well as a transition from analogue to digital, we saw a mono-service to multi-service transition. From 2G to 3G, as well as a mono-media to multimedia transition, we are also seeing transition from person-to-person to person-to-machine interactions, with users accessing video, Internet/Intranet and database feeds. The 3G to 4G transition, supported by such technologies, will see a transition towards a predominance of automated and autonomously initiated machine-to-machine interactions.

Such developments will of course be accompanied by ongoing evolution of already anticipated 3G services such as

- Send/receive e-mail
- Internet browsing (information)
- On-line transaction (e-business)
- Location dependent information
- Access company databases
- Large file transfer.

These services in themselves represent an increase in requirements to access information and for business and commercial transactions, as well as a raft of new location dependent information services, all including significant higher bit rate requirements. Requirements are for a mixture of unicast, multicast and broadcast service delivery with dynamic variation between application services both spacially and temporally. Above all, there is a demand for ease of user access and manipulation, with minimal user involvement - complexity hidden from the user - with intelligence to learn and adapt with use.

From the above it will be seen that 4G will need to be highly dynamic in terms of support for:

- User's traffic
- Air interfaces and terminal types
- Radio environments
- Quality of Service types
- Mobility patterns

In response to this, 4G must itself be dynamic and adaptable in all aspects with build in intelligence. Thus a "software system" rather than a hard and fixed physical system is indicated. Integration is also a key to 4G, needed to reflect the convergence issues already mentioned, and in particular integration of the radio access and the core network elements which must be designed as a whole rather than segmented as in the past. Key drivers to 4G will be:

- Multitude of diverse devices (distributed, embedded, wearable, pervasive).
- Predominance of machine to machine communications
- Location dependent and e-business application
- Extension of IP protocols to mobility and range of QoS.
- Privacy and security
- Dynamic networking and air-interfaces
- Improved coverage mechanisms
- Improved and dynamic spectrum usage.

3. 4G Visions mapping to research topics

The mobile VCE vision for 2010 is embodied in the five key elements as shown in *Fig. 2* They are as follows:

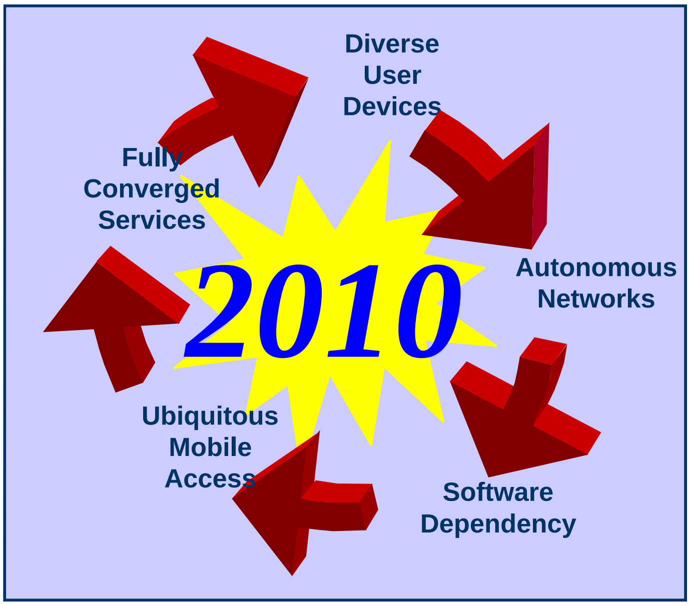

Fig 2: Key elements of 4G vision

Fully converged services
Personal communications, information systems, broadcast and entertainment will have merged into a seamless pool of content available according to the user's requirement. The user will have access to a wider range of services and applications, available conveniently, securely and in a manner reflecting the user's personal preferences.

Ubiquitous mobile access
The dominant mode of access to this pool of content will be mobile, accounting for all voice communications, the majority of high-speed information services, and a significant proportion of broadcast and entertainment. Mobile access to commercial and retail services will be the norm, replacing current practices in most cases.

Diverse user devices
The user will be served by a wide variety of low-cost mobile devices to access content conveniently and seamlessly. These devices will commonly be wearable - in some cases disposable - and will normally be powered independently of the mains. Devices will interact with users in a multi-sensory manner, encompassing not only speech, hearing and sight, but also the other human senses, biological and environmental data pertinent to the application. Special devices tailored for people with disabilities will be commonplace.

Autonomous networks
Underlying these systems will be highly autonomous adaptive networks capable of self-management of their structure to meet users' changing and evolving demands, for both services and capacity. Efficient and cost-effective use of the radio spectrum will be an essential element of their operation, and here too autonomy and self-management will be the norm.

Software dependency
Intelligent Mobile Agents will exist throughout the networks and in user devices, acting continually to simplify tasks and ensure transparency to the user. These Mobile Agents will act at all levels, from managing an individual user's content preferences, to organising and reconfiguring major elements of networks.

Analysis of the underlying technical challenges raised by the above vision and its five elements have produced three research areas, each having a number of distinct research needs. These form the basis of the Mobile VCE Phase 2 research programme. They are by no means exclusive, and have to some extent reflected the research skills available in our academic research teams. These are as follows.

Networks & Services

- Evolved IP protocols for mobile systems
- Ad hoc networking
- Resource management for Multimedia mobile systems
- Mobility management in all IP networks
- Network dimensioning for Multimedia, multirate systems.

Software based systems

- Mobile middleware to support resource management, reconfiguration and service demands
- Software networks and radio
- Software agents
- Distributed mobile management systems
- Integrity and security mechanisms

Wireless access

- Flexible and reconfigurable radio architectures
- Self planning and dynamically reconfigurable access
- Intelligent and co-operative capacity enhancement
- Dynamic spectrum allocation
- Resource metric estimation

- HAPS and satellite integration

4 Research Challenges

In the previous section we have evolved a top level vision which we will now attempt to expand in order to draw out more details of the research challenges that such a vision produces. In doing this we hope at the same time to amplify the visions
It would be unreasonable to expect this vision to remain static - it should evolve and grow with time and research experience. What is given here is a glimpse of the footpath which will lead to the highway of 2010.

4.1 Networks and Services

"To provide multimedia multirate mobile communications anytime and anywhere" - although the aim of 3G can only be partially met. It will be uneconomic to meet this requirement with cellular only. 4G will extend the scenario to an all IP network (access + core) integrating broadcast, cellular, cordless, WLAN, short range systems and fixed wire. The vision is of integration across these network air interfaces and in a variety of radio environments on a common flexible and expandable platform. A "network of networks" with distinctive radio access connected to a seamless IP-based core network.

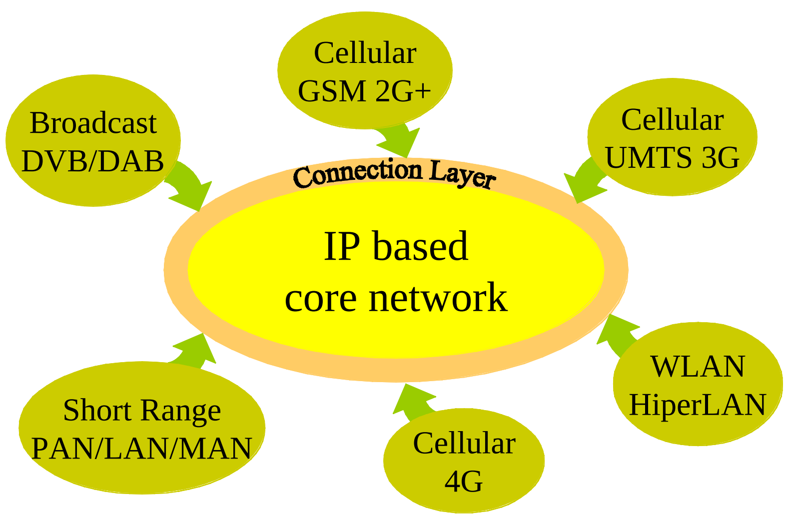

Fig 3:Seamless Connection of Networks

The functions contained in the above will be:
- Connection layer between radio access and IP core including mobility management.
- Internetworking between access schemes - inter and intra system, handover, QoS negotiations, security and mobility.
- Ability to interface with a range of new and existing radio interfaces

A vertical view of this 4G vision (*Fig 4*) shows the layered nature of hierarchical cells which facilitates optimisation for different applications and in different radio environments. In this depiction we need to provide global roaming across all layers.

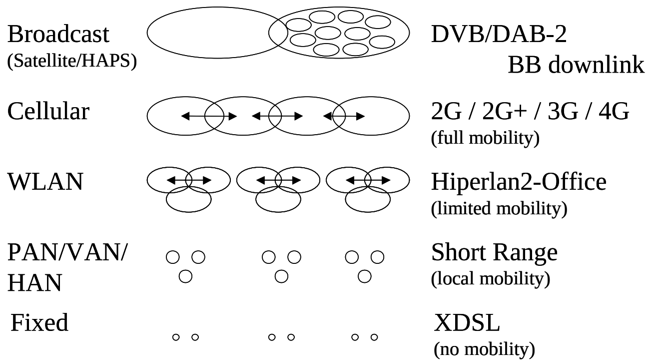

Fig 4: Vertical Hierarchical Networks
[see W.Mohr contribution to (3)]

Both vertical and horizontal handover between different access schemes will be available to provide seamless service and quality of service.

Network reconfigurability is a means of achieving the above scenario. This encompasses terminal reconfigurability that enables the terminal to roam across the different air interfaces by exchanging configuration software (derived from the software radio concept). It also provides dynamic service flexibility and trading of access across the different networks by dynamically optimising network nodes in the end-to-end connection. This involves reconfiguration of protocol stacks, programmability of network nodes and reconfigurability of base stations and terminals.

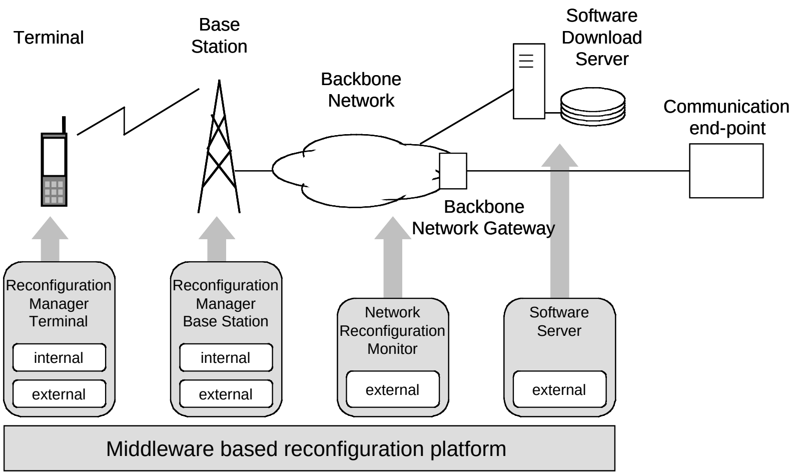

Fig 5: Reconfiguration of Mobile Systems

The requirement is for a distributed reconfiguration control. Fig 5 demonstrates both internal node and external network reconfigurability.

**It is assumed that DVB and DAB will have by 2010 moved on to a new merged standard DVB/DAB - 2*

Internal reconfiguration must control the functionality of the network nodes before, during and after reconfiguration and facilitate compliance to transmission standards and regulation.

External reconfiguration management is required to monitor traffic and to ensure that the means for transport between terminal and network gateway (or other end-points) are synchronised (eg standards etc) and to ensure provision of necessary databases/content servers for downloadable reconfiguration software.

The research challenges are to provide mechanisms to implement internal and external configuration; to define and identify APIs and to design mechanisms to provide compliance of reconfigured network nodes with regulatory standards.

An example of evolved system architectures is a combination of adhoc and cellular topologies. A "mobile ad hoc network" (MANET) is an autonomous system of mobile routers (and connected hosts) connected by wireless links. The routing and hosts are free to move randomly and organise themselves arbitrarily; thus the network wireless topology can change rapidly. Such a network can exist in a stand-alone function or be connected to a larger internet (as shown in Fig.6)

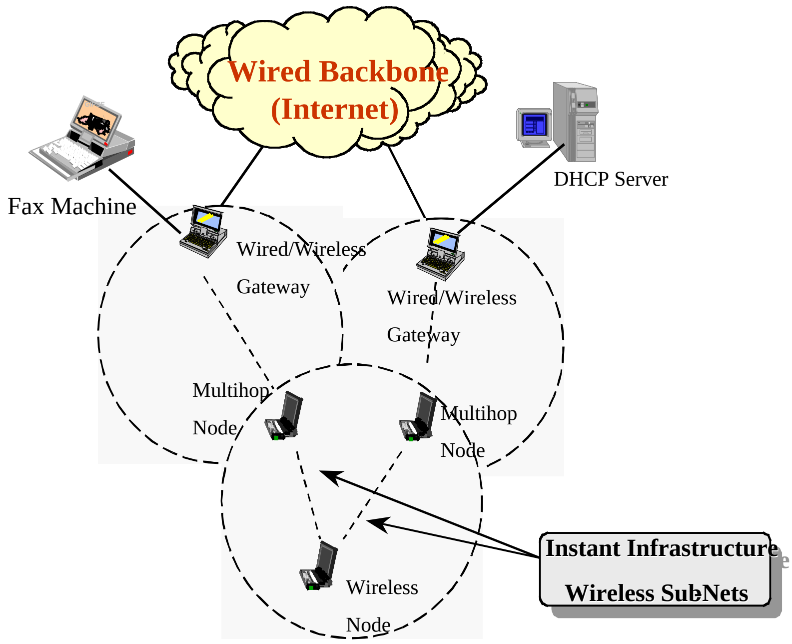

Fig 6: An Integrated Ad Hoc Wireless System

In the current cellular systems which are based on a star-topology, if the base stations are also considered as mobile nodes, the result becomes a "network of mobile nodes" where a base station acts as the gateway in bridging between two remote adhoc networks or gateway to the fixed network. This architecture of hybrid star and adhoc has many benefits such as: self-reconfiguration and adaptability to highly variable mobile characteristics such as; channel conditions, traffic distribution variations, load-balancing, and helps to minimise the mobile location estimation inaccuracies, etc. Together with the benefits we also have some new challenges which mainly reside in the unpredictability of network topology due to mobility of the nodes, this coupled with the local broadcast capability, provides new challenges in designing a communication system on top of an adhoc wireless network.

This will require
- Distributed MAC and dynamic routing support
- Wireless service location protocols
- Wireless dynamic host configuration protocols
- Distributed LAC and QoS-based routing schemes

In Mobile IP networks we cannot provide absolute quality of service guarantees, but various levels of quality can be "guaranteed" at a cost to other resources. As the complexity of the networks and the range of the services increase there is a trade off between resource management costs and quality of service that needs to be optimised. The whole issue of resource management in a mobile IP network is a complex trade-off of signalling, scalability, delay and offered QoS.

As already mentioned, in 4G we will encounter a whole range of new multirate services whose traffic models in isolation and in mixed mode need to be further examined. It is likely that aggregate models will not be sufficient for design and dynamic control of such networks. Effects of traffic scheduling, MAC & CAC and mobility will be needed to devise the necessary dimensioning tools with which to design 4G networks.

4.2 Software Systems

We have already seen in the previous section (4.1) that to effect terminal and network node reconfigurability we need a middleware layer. This consists of network intelligence in the form of object-oriented distributed processing and supporting environments that offers the openness necessary to break down traditional boundaries to interoperability and uniform service provision. The mobile software agent approach as a building block is especially important in offering the ability to cope with the complexities of distributed systems. Such building blocks may reside at one time in the terminal, then in the network, or be composed of other objects that themselves are mobile. Within the mobile system there exists a range of objects whose naming, addressing and location are key new issues. A further step in this development is the application of the Web-service-model rather than the Client/Server principle, late industry tendencies show a shift towards this paradigm and see XML (eXtensible Marup Language) as the technology of the future for Web-based distributed services. However, this technology has yet to prove its scalability and suitability for future application in mobile networks.

In addition to the network utilities there will be a range of applications and services within 4G that also have associated with them objects, interfaces, (APIs) and protocols. It is the entirety of different technologies that underlies the middleware for the new 4G software system.

The "killer application" for 4G is likely to be the personal mobile assistant (PMA) - in effect the software complement to the personal area network - which will organise, share and enhance all of our daily routines and life situations. It will provide a range of functions including:

- Ability to learn from experiences and to build on personal experiences - to have intelligence
- Decision capability to organise with other PMAs and network data bases routine functions - eg. diary, travel, holidays; prompts; (shopping, haircut, theatre, birthdays etc.)
- A range of communication modes - voice, image (superimposing via head-up displays - glasses or retinal overlays etc.) multiparty meetings etc. including live action video of us and our current environment
- Navigation and positioning information and thus location dependent services
  - Location detection and reporting for children, pets and objects of any sort
  - Vehicle positioning and route planning, auto pilot and pedestrian warnings
  - Automatic reporting of accidents (to insurance co., rescue services and car dealers)
- Knowledge via intelligent browsing of the internet
- e-business facilities - purchasing and payment
- health monitoring and warnings
- Infotainment - music, video and maybe VR

Of course the key to all this is "mobility" - we need to have the "PMA" whenever and wherever we are, and this places additional complexity on network and service objects and the agents that process them.

Specifically we need to consider what are the metrics that determine which objects follow the user. Some objects can move anywhere; others can move in some directions or within a constrained area. If they can move, how will the existing service determine if resources are available to support them in their new (temporary) home? Will they still be able to function? What kind of computing architecture and middleware platforms will be capable of supporting thousands; perhaps millions, of such objects?

Aspects of security pervade the whole of this area. Rules of authentication, confidentiality and scalability and availability must now be applied to objects that are continuously mobile. A whole set of valid conditions at one time and place may be invalid if transferred to another. Integrity and correctness issues apply when mechanisms that support applications are used in practice in the presence of other distributed algorithms. Issues such as liveness, safety and boundedness, consistency, isolation and durability - execution semantics need to be evidenced for extension to the mobile environment.

Distributed management tools, in a complimentary way, will allow a certain level of monitoring (including collection of data for analysis) control and troubleshooting. The management tools currently available do not encompass mobility efficiently and hence this is another important area of research.

The aim of the research in this area is to develop tools that can be used in 4G software systems. Whence specific scenarios are being addressed in order to focus the issues as follows:

- e-commerce – including microtransactions, share trading and business internal transactions.
- Home services – ranging from terminal enhancements (e.g. enhancing the display capabilities by using the TV screen as display unit for the terminal) to security systems and house keeping tasks.
- Transportation systems – itinerary support, ticketing and location services are to be targeted in this area.
- Infotainment on the move – will show needs for software and terminal reconfiguration and media-adaptation.
- Telemedicine and assistance services – emergency team support, remote/virtual operations, surveillance of heart patients etc. are possible stages for this scenario.

This list of scenarios can arbitrarily be expanded, also into non-consumer areas (i.e. military and emergency services), however the preconditions for service delivery and demands on the network infrastructure remain the same: they will have to be adaptable to meet the 2010 current user-requirements. Support for these scenarios may be given by intelligent agents, which may represent the terminal within the network to manage the adaptations or customisations of the communication path. On an application or service layer they may be used in addition to complete business transactions for the user (e.g. booking a theatre ticket/flight/etc.) or to support other services. Furthermore, distributed software entities (including the variety of models: from objects via agents to the web-service model) will encompass management and support for applications and services as well as for user *and* terminal mobility.

4.3 Wireless Access

In the previous two sections we have looked at the type of network and the software platforms needed to reconfigure, adapt, manage and control a diversity of multimedia, multirate services and network connections. We have seen that there will be a range of radio access air interfaces optimised to the environments and the service sets that they support. The reconfigurability and the middleware flow through to the wireless access network. The radio part of the 4G will be driven by the different radio environments, the spectrum constraints and the requirement to operate at varying and much higher bit rates and in a packet mode. Thus the drivers are:

- Adaptive reconfigurability - algorithms
- Spectral efficiency - air interface design and allocation of bandwidth
- Environment coverage - all pervasive
- Software - for the radio and the network access
- Technology - embedded/wearable/low power/high communication time/ displays

It has been decided within MVCE not to become involved in technology issues or in the design of terminals. This is a large area in itself, which is much closer to products and better suited to industry. The remaining drivers are all considered within the research programme.

Using adaptive signal processing jointly at both the base station and the mobile, it is possible, in principle, to significantly increase the effective bit rate capacity of a given bandwidth. Within 3G this has been limited to processing at the base station and so the challenge is to migrate this to the terminal, and most importantly, to make the two ends cooperative. Such techniques require close co-operation between the base and mobile stations in signalling channel quality information, whilst making decisions and allocating resources dynamically. In addition, the capabilities of both ends of the link must be known reciprocally as the channel varies in both time and space. In order to continuously optimise a link, the wireless network must acquire and process accurate knowledge of metrics which indicate the current system performance. eg. noise, inter and intra system interference, location, movement variations, channel quality prediction. Such information and its accuracy must be passed to higher layers of the system protocol which will make decisions and effect resource allocation. The emphasis on the base station in 3G is obvious as this has the resource, real estate and capacity to implement spatial-temporal processing involving dsp for antenna array processing together with advanced receivers architectures. The challenge will be to migrate this to the much smaller terminal via efficient electronics and algorithms that will still allow a range of services and good call time. The availability of individual link metrics can also be used at a network level to dynamically optimise the network radio resources and produce a self-planning network.

Arguably the most significant driver in the wireless access is the bandwidth availability and usage and whereabouts in the spectrum it will fall. Currently 3G is based around 2GHz but with limited spectrum available even with the addition of the expansion bands. The higher bit rates envisaged for 4G will require more bandwidth - where is this to be found? The scope for a world wide bandwidth allocation is severely constrained and, even if feasible, will be very limited. The requirements are thus for much more efficient utilisation of the spectrum - "spectrum efficiency", and perhaps new ideas for system co-existence. Taking spectral efficiency within a fixed bandwidth we need to seek a more efficient air-interface - this involves multiple access, modulation, coding, equalisation/interference cancellation, power control etc. Coupling with previous comments it is clear that all components of this air-interface must be dynamically adaptive. As the whole network is to be IP based this will mean extremely rapid adaptation on a burst basis. Within 4G we need to accomplish this at much higher and variable bit rates as well as in different environments (indoor, outdoor, broadcast etc.) and in the presence of other adaptive parameters in the air-interface. In time-domain systems equalisers would need to be adaptive, raising questions of complexity. For CDMA, systems could use multicodes and adaptive interference cancellation which again raises complexity issues. Alternatively one could move to OFDM-like systems (as in WLANs) which offer some reduction in complexity by operating in the frequency domain but raise other issues such as synchronisation. The choice of the multiple access and the adaptive components of the air-interface will need to be made based upon the ease of adaptation, reconfigurability and on complexity. Significant research challenges exist in this area of flexible advanced terminal architectures which are not rooted solely in physical layer problems.

A further aspect of spectrum efficiency relates to the way in which regulators allocate bandwidth. The current practice of exclusive licensing of a block of spectrum is arguably not the most efficient. It would be much more efficient to allow different operators and radio standards to co-exist in the same spectrum by dynamically allocating spectrum as loading demands. Indeed the higher bit rate services may need to spread their requirements across several segments of spectrum. There would then be a need for a set of rules to govern the dynamic allocation of the spectrum - a self organising set of systems to maximise the use of spectrum and balance the load. Given the degree of co-operation and the processing already envisioned this should be a realistic aim.

A great deal of work on the characterisation of radio environments has already been performed in the 2GHz and 5GHz bands within the first phase of Mobile VCE's research and spatial-temporal channel models have been produced. However, 4G systems will incorporate
smart antennas at both ends of the radio link with the aim of using antenna diversity in the task of cancelling out interference and assisting in signal extraction. This implies that direction of arrival information, including all multipath components, will be an important parameter in determining the performance of array processing techniques. There is a need to augment models with such data for both the base station and the terminal station. A more open question is the determination of where to position the next frequency bands for mobile communications. An early study is needed here in advance of more detailed radio environment characterisations.

Coverage is likely to remain a problem throughout 3G, and with the network of networks structure of 4G, together with the addition of multimedia, multirate services it will continue to present challenges. We have already seen that the likely structure will be based upon hierarchical - (macro, micro and pico) cells. Superimposed on this is the mega cell which will provide the integration of broadcast services in a wider sense. Until now, it has been assumed that satellites would provide such an overlay and indeed they will in some areas of the world. However, another attractive alternative could be high altitude platform stations (HAPS) which have many benefits particularly in aiding integration.

HAPS are not an alternative to satellite communications but more a complementary element to terrestrial network architectures mainly providing overlaid macro/micro cells for underlayed pico-cells supported through ground-based terrestrial mobile systems. These platforms can be made quasi-stationary at an altitude around 21-25 km in the stratospheric layer and project hundreds of cells over Metropolitan areas.

Due to the large coverage provided by each platform, they are highly suitable for providing local broadcasting services. A communication payload supporting 3G/4G and terrestrial DAB/DVD air-interfaces and spectrum can also support broadband and very asymmetric services more efficiently than 3G/4G or DAB/DVB air-interfaces could individually. ITU-R has already recognised the use of HAPS as high base stations as an option for part of the terrestrial delivery of IMT-2000, in the Bands 1885-1980 MHz, 2010-2025 MHz and 2110-2170 MHz in Regions 1 and 3 and 1885-1980 MHz and 2110-2160 MHz in Region 2. [Recommendation ITU-R M. (IMT-HAPS)].

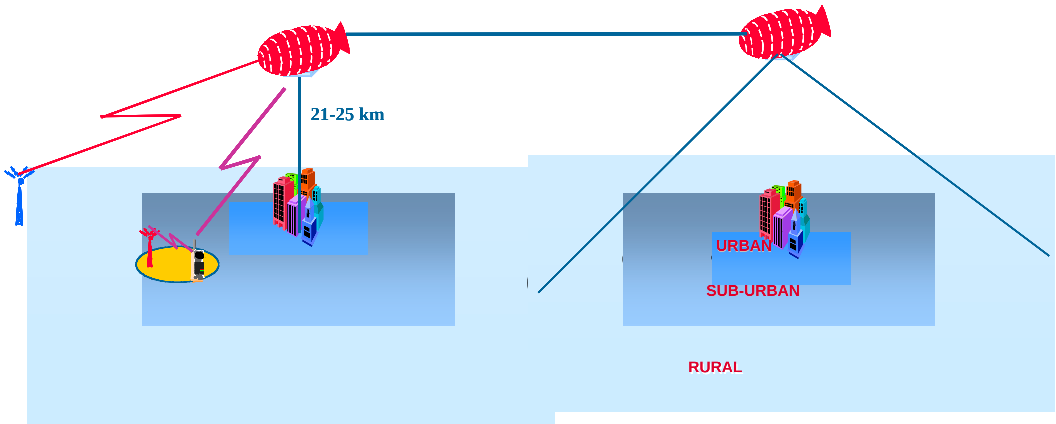

Fig 7: HAPS providing integrated coverage

HAPS have many other advantages in reducing terrestrial real-estate problems, achieving rapid roll-out, improved interface management to hundreds of cells, spectrally efficient delivery of multicast/broadcast and provision of location based services and of course integration. The research challenge is to integrate terrestrial and HAPS radio access so as to enhance spectral efficiency and preserve QoS to the range of services offered.

Software, algorithms and technology are the keys to the wireless access sector. Interplay between them will be the key to the eventual system selection, but the Mobile VCE's research programme will not be constrained in this way. The aim is to research new techniques which themselves will form the building blocks of 4G.

5 Conclusions

It is always dangerous to predict too far ahead in a fast moving field such as mobile communications. Almost by definition the eventual 2010 scene will not match exactly that depicted in the 4G vision herein. However, we feel that the key elements, fully converged services, ubiquitous mobile access, diverse user devices, autonomous networks and software dependency will persist. Their exact technological manifestation in the eventual 4G may depart from our initial ideas outlined in Section (4) Research Challenges, but the important thing is that we have created a scenario that points us in the right direction. Our 4G vision has been based upon both migration from 3G and from drivers and deficiencies of the latter system as well as perceived user trends. We have deliberately herein steered clear of "user scenarios" which have been well researched by others (see 1-2). The 4G vision is thus a technology vision, but it has been well informed by user trends and by such user visions as mentioned previously. We have also participated in an EU Visions programme (WSI)[3] which may be consulted for a wider perspective.

The 4G Vision is a living document which we intend to update and amend as time and knowledge progress. It will act as the umbrella vision to a large research programme and place in context the detailed research work that will take place in the various areas. In this respect it will help to continuously steer the research as it progresses and we hope therefore to make it more relevant and beneficial.

6 References

1 Irvine J. et al - Mobile VCE Scenarios - document produced as part of the Software Based Systems work area within the Mobile VCE Core 2 research programme, September 2000. http://www.mobilevce.com

2 "Visions" - Foresight ITEC Group. http://www.foresight.gov.uk

3 Wireless Strategic Initiative - EU project. http://www.ist-wsi.org

Glossary

ANSI	American National Standards Institute
API	Application Programming Interface
BRAN	Broadband Radio Access Network (ETSI)
CAC	Connection Admissions Control
CDMA	Code Division Multiple Access
DAB	Digital Audio Broadcast (standard)
DVB	Digital Video Broadcast (standard)
EDGE	Enhanced Data for GSM Environment
ETSI	European Telecommunications Standards Institute
FDD	Frequency Division Duplex
GPRS	General Packet Radio System
GSM	Global System for Mobile Communications
HAN	Home Area Network
HAPS	High Altitude Platform Stations
HIPERLAN	High Performance Local Area Network (ETSI standard)
HTML	Hypertext Markup Language
ITU	International Telecommunications Union
IMT-2000	International Mobile Telecommunications 2000 (3G standard ITU)
MAC	Medium Access Control
MAP	Mobile Applications Protocol
OFDM	Orthogonal Frequency Division Multiplex
PAN	Personal Access Network
SMS	Short Message Service
UMTS	Universal Mobile Telecom System (3G standards ETSI)
UTRA	UMTS Terrestrial Radio Access
VAN	Vehicle Access Network
VHE	Virtual Home Environment
WAP	Wireless Access Protocol
WCDMA	Wideband Code Division Multiple Access
WLAN	Wireless Local Area Network
WML	Wireless Markup Language
XDSL	X (various) Digital Subscribers Line