Parallel Transmission Frequency Division Multiplexing

The essence of the parallel transmission is to use the special ADSL channel characteristics by transmitting multimedia layers simultaneously, each occupying a set of subchannels. As shown in Figure 6.52, frequency division multiplexing the data streams corresponds to different multimedia layers. The three service layers are transmitted in the same slot, each occupying a set of subchannels. Although the time slots are still grouped into frames, for every time slot in a frame, the power and bit-rate allocation remain the same. In general, data from the important layers are transmitted through the subchannels with better channel performance, that is, larger channel gain, lower noise variance or simply the good subchannels. Such assignment provides reliable transmission to the most important layers without large power consumption. In contrast to the serial transmission, the parallel transmission can integrate various traffic flows with different QoS requirements without any frequent channel parameter changes.

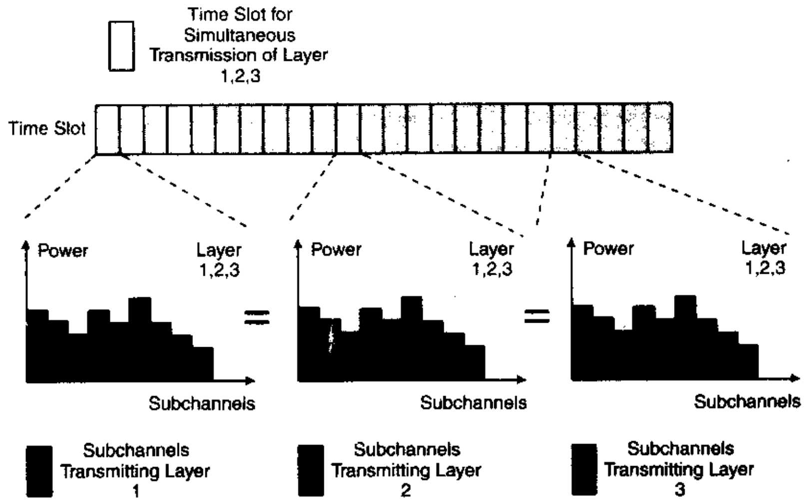

Figure 6.52 Parallel transmission for multimedia data over ADSL [6.145]. ©2000 IEEE.

The error-performance distribution of the parallel transmission is quite different from that of the serial transmission, as can be seen from Figure 6.53. It achieves constant error performance across the time slots, but different error performances at the subchannels transmitting different layers.

6.6 Internet Access Networks

In building an infrastructure for delivering all the services to the consumer, the most critical part is the links to the subscriber's home. This so-called last mile has some unique characteristics. It generally consists of copper wire pairs over which analog voice signals travel to a telco central office and is therefore limited in its information-carrying capacity. Replacing copper pairs with

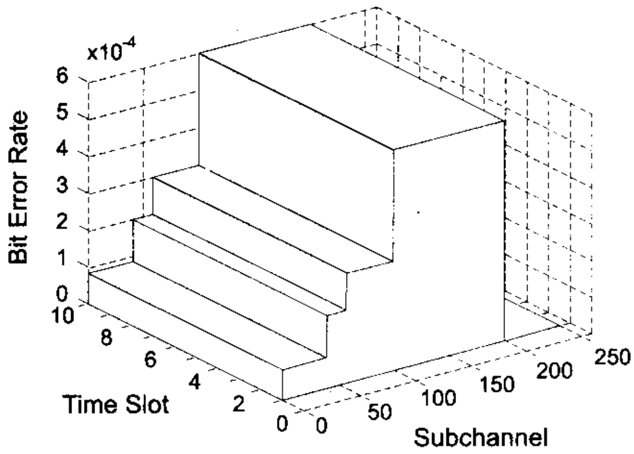

Figure 6.53 Error performance across the time slots and subchannels for parallel multimedia data transmission across ADSL [6.145]. ©2000 IEEE.

optical fibers would be a good way to speed Internet access, but, in fact, the amount of revenue foreseen from each residential and small business customer cannot economically justify such a move. From the sociopolitical point of view, faster Internet access for everyone is considered important. Out of about 120 million United States households, 40% have at least one computer. Of the computerized households, some 60% access the Internet regularly. If it is assumed that 10% of those regular Internet accesses become early adopters of fast Internet access services, 2.88 million households will be involved, each of which may be expected to spend about $40 a month on the second phone line and Internet access. Accordingly, the market for combined services offering a second telephone line only for Internet access over a short term of one year or so may be conservatively estimated at $1.3 billion a year. This is an impressive amount, and it is just the beginning.

ADSL offers asymmetric rates of transfer of data to and from the Internet. The uplink rates can go up to 768 Kb/s and downlink rates are 6 to 8 Mb/s, depending on the length and condition of the local loop, or the wiring between the customer's premises and the telco central office.

Cable companies bring analog TV signals across optical filters to their neighborhood distribution points (headends), where the signals are distributed to residences by coaxial cables. The combination of fiber and coaxial cable, which can carry high-speed data as well as TV signals, is known as hybrid fiber coax (HFC). Each distribution point typically serves 200 to 500 residences. The extent of the network of a cable TV operator is measured in terms of homes passed, that is, the number of homes adjacent to which the operator's cable passes, regardless of whether those homes have been signed up as customers. Realistically, cable modems are capable of passing data upstream at speeds of 200 Kb/s to 2 Mb/s and downstream at speeds up to about 10 Mb/s.

Cable modems, capable of operating at higher speeds than ADSL, have some serious drawbacks. The cable link to a residence is shared among many users, so if many of them decide to log on to the Internet at the same time, achievable communication speeds may plunge.

Because the lines are shared, a hacker may be able to drop on a neighbor's connection to the Internet or on an intranet, which is a security problem that may be serious to some users. Consequently, a customer who happens to be a road warrior will be unable get access into the Internet at airports or hotels through his laptop computer at his usual data rate. If he is able to connect at all, it will be through a dial-up modem at a much lower speed.

The most valuable benefit of wireless services is that they make access possible for people who are on the move. They are also attractive in certain cases where the user is stationary. Digital cellular telephones are quickly becoming the main communication tools for people on the move. Although they are good for retrieving email and checking stock quotes, the R&D is targeting multimedia communications. Multimedia communications are available, but at present only to stationary users. Satellite broadcasts, for example, allow fast download of Internet contents with a return path (that is, the uplink from the user computer to an Internet service provider) on a dial-up modem. For a rural user, it is possible to use Local Multipoint Distribution Services (LMDS). This option uses millimeter-wave radio at frequencies of about 30 GHz. A typical installation has a central base station with an unidirectional antenna serving many residences, each of which has a directional dish aimed at the base station. The service is theoretically capable of sustaining a data transfer rate of about 30 Mb/s. The systems work well provided that the users are within about 3.5 Km of the local transceivers. Yet another viable alternative is to access the Internet across the unlicensed 2.4 GHz band. With this approach, the service provider broadcasts Internet data, using digital spread-spectrum wireless technology from an antenna on a tower. The range is some 60 Km. Computer manufacturers and software developers also have legitimate interests in both methods of fast Internet access: ADSL and cable modems. Faster means of Internet access generate demand for faster computers and newer software.

6.6.1 DSL Networks

Whether by phone or by mail, local phone companies will soon be urging their customers to sign up for an Internet service speedier than anything achievable with today's dial-up modem. A technology known generally as DSL is responsible. It needs only a single twisted wire pair to provide both Internet access and conventional analog telephony. In the future, it may also be used to deliver pay-per-view video. Most of these services will require minor changes to the phone wiring in subscribers' homes [6.146]. The most suitable version of the technology for residential broadband access to the Internet is generally held to be ADSL. Its downstream bandwidth, from the Internet to the home, may reach 6.144 Mb/s and its upstream bandwidth, from the home PC to the Internet, may reach 640 Kb/s [6.147]. The asymmetry in ADSL transmission speeds matches the flow of data to and from the Internet. In a typical Internet session, after all, a Web surfer sends short messages upstream to request data and is bombarded with information in return. It is this downstream transmission rate that limits the usefulness of most connections [6.148].

From a performance point of view, the big difference between DSL technologies and cable modems is that DSL gives each customer a dedicated link to the central office of the local phone company. On the other hand, with cable modems, several users share a single coaxial cable [6.149].

Cable TV operators, ever since they saw the promise of their coaxial cables, have been busily equipping homes with high-speed access to the Internet. Their coax links penetrate more than 90% of residences in the United States, with European and Japanese percentages not far behind and with rapid growth in the nonindustrial world as well. Telephone companies were slower off the mark, having made impressive earlier attempts to capture the residential Internet access market with their ISDN. The sudden threat posed by cable TV companies is now forcing them to deploy ADSL.

ADSL does have its problems, however. Loading coils are one. Because of the distributed capacitance and resistance along their lengths, phone wires attenuate and distort voice signals, with effects that increase with frequency and distance. To equalize the lines across the frequency range of interest for voice communication (up to 4 KHz), phone companies install inductors in their longer lines. These loading coils improve the frequency response within the voice band, but at a price: They increase attenuation for signals above that band. Consequently, they must be removed for high-speed data transmission.

Signal dispersion is another problem with high-frequency signals. The physical characteristics of transmission lines are such that signals of different frequencies propagate at different velocities. Pulses, which represent data and are made up of several frequencies, tend to spread out as they propagate down a line, eventually overlapping with each other. This effect is known as intersymbol interference and limits the data rate that can be supported. Like attenuation, the effects of dispersion get worse with frequency and line length. Both near-end and far-end cross-talk are also problems. The first arises when a receiver is located at the same end of a cable as a transmitter operating in the same frequency band.

The frequency band used for full-rate ADSL is broken into three parts as shown in Figure 6.54. It can be seen that in the FDM used for ADSLs, the frequency band from DC to 1.1 MHz is divided into three subbands (top). The first is used for analog voice. Namely, the 0 to 4 KHz range is reserved for voice telephony; the portion between 25 KHz and 138 KHz for upstream data to the Internet and the rest of the band, up to 1.1 MHz, for downstream data from the Internet to the PCs. If some form of echo cancellation technology is used, the downstream bandwidth may be expanded. Echoes are signals generated by the local transmitter that get fed to the local receiver due to coupling between wires. An echo canceller takes care of echoes because it knows what was transmitted and can subtract it from what was received. Basically, two types of ADSL modems now coexist: Carrierless Amplitude Phase (CAP) and Discrete Multitone (DMT). They differ in how they perform line coding, that is, in how they modulate digital data onto an analog carrier. CAP uses quadrature amplitude where a pair of equal-frequency signals are varied to create between 4 and 1,024 discrete line conditions or symbols. Each symbol represents several bits, and the actual number is dependent on the total number of possibilities.

Example 6.2 In a CAP-4 scheme, with a total of four possible symbols, each symbol represents two bits. A phase of 0 degree could represent 00, 90 degrees could be 01, 180 degrees

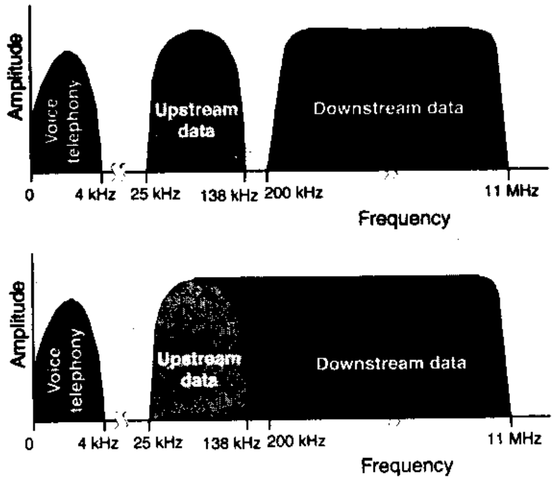

Figure 6.54 The frequency band used for full-rate ADSL [6.150]. ©1999 IEEE.

would be 10 and 270 degrees would be 11. In a more practical example, CAP-16 uses 12 different phases and then adds a further four symbols by repeating four of the phases, but at half the amplitude. In this scheme, then, a total of 16 distinct combinations of bit patterns, starting from 0000 and ending with 1111, can be represented. Because each symbol represents 4 bits, CAP-16 can operate at a line speed of, say 10000 symbols per second, and yet can transmit data at 40 Kb/s.

The business of providing the physical medium for the transport of data to the Internet is to be different from the business of providing Internet services. The first is called a network service provider. Such a provider would deploy the ADSL technology offering users a fast access path to the Internet. After the path is established, the ISP's home page, also known as a portal, is the first to come up on the screen of the user's computer.

6.6.2 Cable Access Networks

Today's cable plants suit one-way communication, from the headend outward, and so are perfectly adapted to broadcast television. For Internet access or telephony, however, traffic must flow in both directions, so a range of frequencies between 5 and 42 MHz is allocated for upstream signals, both analog and digital. Any discussion of access to the Internet must emphasize that, for most residential uses, other than video or desktop conferencing, the exchange of data is asymmetrical. The request an Internet user sends to look up a Web site or search for some information consists of very few packets but may trigger a deluge of data in the other direction. This deluge is one of the main problems that cable modems aim to solve.

Because the connection between a cable modem and the Internet is packet switched, the link is always active. As soon as the computer is turned on, it begins a new session with the Internet, and the ISP dynamically assigns it an IP number, which remains valid until the session is terminated. A large provider with thousands of clients may assign local IP addresses to clients using its proxy server to interface with the outside world, that is, outsiders will see the proxy server's address, instead of the local address, and all communication between a user and the outside world will take place through the proxy server. The telephone service provided by the cable operators is also always on in that it is not interrupted by power outages. Should utility power fail, backup batteries at the cable system's nodes would keep the phones working even when other services go down. This feature is akin to the lifeline service provided by phone companies.

For some years now, multiple-system operators have tempted their clients with Internet and cable TV service. This they did with a splitter box and a cable modem of proprietary design. The devices cost a lot and have only a single source. Hence, there arose a strong desire for a set of specifications that would be widely acceptable to the industry.

Of concern in sending digitized voice across a packetized system is that delay and delay variations (jitter) be kept within fairly tight limits. Small packets are desirable because they keep the delay down, but are undesirable because they increase overhead. Long packets have lower overhead, but take longer to fill. In a multiuse hybrid fiber-coax system, a packet size of an IEEE 802.3/Ethernet frame (1,518 bytes, maximum) would represent a compromise among many requirements.

The question often arises as to what mechanism should be adopted for the MAC layer. Downstream data present no problem as it is broadcast to all subscribers alike. However, upstream data travels from all subscribers to just one headend, risking collisions between packets. In a conventional carrier-coax system, because cable modems do not transmit and receive in the same frequency band, a station cannot sense collisions between its upstream packet and packets sent from other stations. A Time Division Multiple Access (TDMA) method has been chosen to coordinate the upstream hybrid fiber-coax transmissions. In implementing this, the headend regularly broadcasts a message to see if any station that so wishes undergoes an initialization procedure that synchronizes its clock with that of the headend and determines how long it takes for a signal to traverse the path connecting them.

Electrical noise in the upstream direction can give both end-users and the multiple-systems operator a big headache. Noise can seep into the system from many sources, including home appliances. The use of cheap cabling or a poorly designed splitter can make the problem worse. In a hybrid fiber-coax configuration, each drop cable and neighborhood feeder cable acts as spokes of a sprawling antenna, and the headend acts as a giant receiver and accumulates all their noise. The collected noise degrades all spokes, thus worsening the situation. There are various ways to minimize the ingress of spurious noise. The first is to protect the service area thoroughly for potential sources of their effects. The second is to use, as far as possible, the cleanest part of the 5 to 42 MHz band, which is the portion that is between 21 MHz and 27 MHz. The third is to reduce the number of households served by each headend or regional hub. Admittedly, reducing the number of hubs in a given area will raise system costs. However, in the long run, the investment is worthwhile for an operator committed to good customer service.

The cable modem and terminal equipment manufacturer, Com21, has taken an interesting approach to tackling the noise in the upstream, or return, paths. The key is the company's eight-port card, known as the Return Path Multiplexer (RPM). Located at the fiber node, the eight-port card multiplexes eight return paths so that the ComComtroller, at the headend, can schedule such that, at any given instant, it receives data from only one of those eight upstream links. The seven other return paths are blocked so long as one is open. Thus, at no time can the noise from eight return paths accumulates at the headend or hub, obviating the expense of deploying the extra hubs or headends otherwise necessary to reduce the number of subscribers and the noise per hub.

The rapidity of the progress being made in cable modems may seem to imply that every important problem has been solved, but many remain. Some customers complain about low speed, which may be caused by noise, and costs are always an issue.

Of course, there is one additional aspect to accessing the Internet, the selection of the ISP. By now, many surfers have developed preferences for or aversions to particular providers. However, cable companies do not appear to be willing to open their networks to outside service providers, and, unlike telephone companies, they are not required to do so.

6.6.3 Fixed Wireless Routed for Internet Access

Wireless technology is useful even in congested urban areas. For wireless links, construction and equipment costs have a ratio of roughly 20:80, whereas, for a terrestrial optical-fiber link, the ratio would be more than reversed, about 90:10. Thanks to that enormous cost advantage, wireless systems can be a boon in nonindustrial countries with little telephone infrastructure. Further, the systems may be expanded and scaled up incrementally as groups of new customers sign up, a strategy that demands less initial investment [6.151].

In industrial countries, wireless local loops make sense in many situations. Wireless links are increasingly being used for broadband services. A fixed wireless provider often operates in what is referred to as the point-to-multipoint mode. Its antenna communicates with several different clients' antennas installed within a well-defined region. An ISP might use a point-to-point link to connect its hubs to a distant point that, in turn, is connected to the Internet backbone across a high-speed wired link. A radio link that supports analog voice telephony acts as a simple local loop between a user and a telephone company's central office. However, there is another way to use fixed wireless technology to make voice calls, the Internet. In such case, the user's computer would digitize the voice and set up an IP address with the provider, for example, communication between the user and the Internet probably with Ethernet frames. At the provider's facility, the packetized signals would be converted into conventional phone signals and then to the PSTN. For large business customers or for multiple dwelling units, a competitive local exchange carrier might provide a broadband wireless connection by way of a private branch exchange at the customer premises. The competitive local carrier would then provide the gateway to the public phone network [6.152].

In the United States, the Federal Communications Commission (FCC) has set aside 15 frequency bands for use in commercial fixed wireless service at frequencies of 2 to 40 GHz. In other countries, frequencies are also allocated by the national telecommunications regulating authorities. Several frequency bands are available for fixed wireless services. The 2.4000 to 2.4835 GHz band is popular with many operators because it is unlicensed and no fee is required for its use. It is also popular with equipment vendors because it is available internationally and equipment made for it has a relatively large market.

A technology for wireless services does exist in the form of spread-spectrum techniques, and some licenses were assigned a number of years ago for wireless services [6.153]. The licensees are referred to as the band's primary communications user. With time, the FCC was persuaded to allow other operators—any who so desired, in fact—to use this band as well. The band was then declared to be unlicensed, but not totally regulated. Lower than power levels specified by the FCC, newer operators must use spread-spectrum technology so as not to interfere with the primary users. The primary users did not have to switch the spread spectrum and are allowed higher power. It is possible that more than one operator in this band will use the same scheme for spread spectrum in the dense geographical area. Hence, the likelihood of one interfering with the other is small. However, the possibility of interference exists, which is the price one pays for using an unlicensed frequency.

Two more unlicensed Industrial, Scientific, and Medical (ISM) bands in the United States are over the bands spanning 5.725 to 5.875 GHz and 24.0 to 24.25 GHz. The former is also known as Unlicensed National Information Infrastructure (UNII) band. In the United States, the rest of the fixed wireless spectrum is taken up by licensed bands. Probably the least used bands so far for Internet access are those that cover 2.1500 to 2.1620, 2.5960 to 2.6440 and 2.6500 to 2.6800 GHz, providing 13 channels of 6 MHz each. The first two were licensed back in 1970 when they were called Multipoint Distribution Services (MDSs), to broadcast 6 MHz television channels. In 1996, the FCC expanded the band to cover its present range and allowed for multichannel services called Multichannel Multipoint Distribution Services (MMDS).

At present, the market for local multipoint distribution systems is expanding quite rapidly and is expected to exceed $2 billion by 2003. After all, the technology has the bandwidth needed for broadband services, and its recurring cost is modest. It is therefore expected to appeal to a broad range of customers.

From a service provider's point of view, large enterprises would be the most profitable customers in terms of profit per invested dollar. Which system best meets the needs of a user organization has many determinants, including the number of customers expected in the near and long term, the estimated number of users logged on to the Internet at any one time, the purchase price, the cost of running the system and the nature of the terrain across which the service is to be deployed. Perhaps the most important factor to a service provider is the cost of license for using a piece of the spectrum, unless it is in the unlicensed ISM band. The wireless communications market hums with news about purchases of licenses and acquisition of companies handling licenses. The design of the system is critically dependent on the frequency used and the transmitted power. In general, higher the frequency means the more complex that the system is.

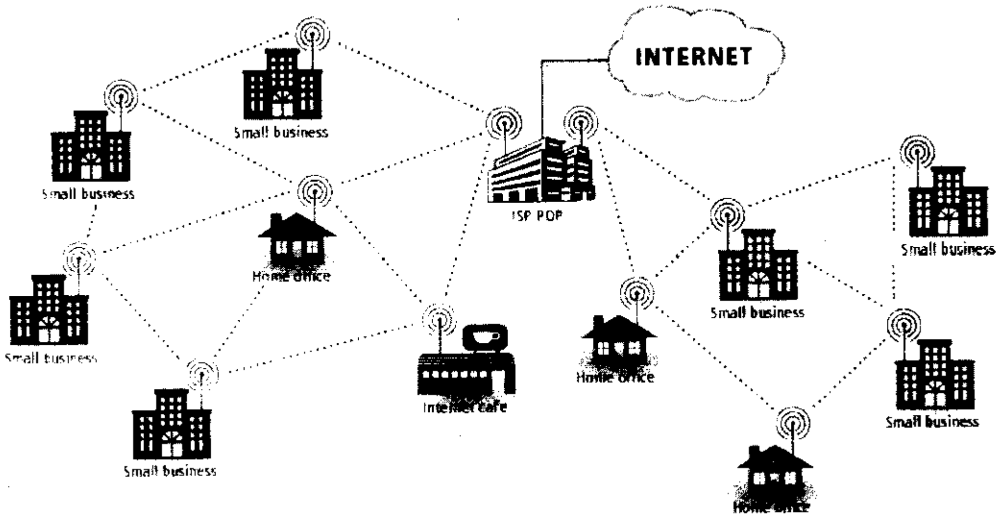

Figure 6.55 Multipoint-to-multipoint wireless scheme.

In a multipoint-to-multipoint wireless environment, multiple logical links exist between one receive/transmit point and its neighboring points in the system (Figure 6.55). Each link between two points may also have different characteristics, such as transmit power, data rate and reliability. All these factors called for a new approach toward the physical media access and network protocols. The next generation architecture attempts to formalize these new protocol modules, the function of each one of them and how information could be exchanged between the modules. The experience gained in the GloMo's Wings project enabled Rooftop to batch its own commercial Internet radio powered by Internet Radio Operating System (IROS) operated in the unlicensed 2.4 GHz IMS band. In a Rooftop Spirit environment, all nodes have same transmit/receive capabilities. However, at least one node is wired to the Internet backbone by a high-speed access line. All of the nodes in the system are aware of all of the other nodes and relay the information forward and backward from node to node. Rooftop calls this node an Airhead. A single Airhead can support 10 to 15 clients and can maintain a reasonable rate of data flow. IROS ensures that all the nodes in the system are aware of all the other nodes and relay information from one to another, as needed, to get packets to their intended destination. One end of every path will invariably be the ISP's point of presence (POP).

6.7 Multimedia Across Wireless

The explosion of technological advancements and the success of the second-generation digital cellular systems (for example, Global System for Mobile (GSM) and Personal Digital Cellular (PDC) have established wireless communications as indispensable in modern life. Because of the low-cost and low-consumption characteristics of emerging wireless products targeted at low- to medium-bit-rate services, these products are expected to play an important role in wireless communications in the next few years. Wireless in multimedia communications (for example, audiovisual telephony and videoconferencing) require medium- to high-bit-rate channels (64 Kb/s to 2Mb/s per user). Therefore, for these applications, it will be necessary to have broadband wireless networks that support bit rates in excess of 2 Mb/s per radio channel, where each radio channel could be shared by multiple users or sessions. In addition, these services have to be provided with some QoS guarantees across their respective, error-prone wireless connections. In order to achieve these goals, one has to address the following key issues:

- How to increase the capacity of wireless channels
- How to provide QoS in a cost-effective way
- How to combat the wireless channel impairments

These questions and related technical issues have been addressed in numerous overview and research papers in the field. For example, an overview of Personal Communication Systems (PCSs) that can provide timely exchange of multimedia information with anyone, anywhere, at any time and at low cost through portable handsets is given in Klee and Oui [6.154]. An overview of various multiple-access schemes is provided. In Hanzo [6.155], various wireless multimedia concepts are described, together with sampling and coding theory, cellular concepts, multiple access, modulation and channel-coding techniques. An overview of various emerging wireless broadband networks in Europe as well as a discussion about the frequency spectrum issue are presented in detail in Mikkonen et al. [6.156]. An overview of the status of wideband wireless local access technologies can be found in Palavan et al. [6.157]. Reference 6.158 gives an overview of wireless broadband communications by addressing some of the applications and services that are foreseen as well as some of the technical challenges that need to be solved.

In addition to the continuing interest in wireless audiovisual communication applications mentioned, a great deal of interest has emerged in higher-end wireless multimedia services. However, current wireless networks, which are primarily low-bit-rate narrow-band systems targeted for voice or data, are inadequate for supporting audiovisual communication applications or high-end multimedia services.

Most multimedia services tend to be real-time in nature, that is, the data being transported need to get to the destination by a certain time in order to be useful. This implies the need to develop techniques for call admission, bandwidth allocation and the handling of real-time variable rate streams. These are problems that apply to wired networks as well and are not, therefore, unique to wireless multimedia communication systems.

The two major protocol-related problems in wireless multimedia concern medium access and QoS. Wireless systems are inherently multiple medium access in nature and therefore need to have a reliable MAC layer that also supports QoS [6.159].

Audio, video and graphics need to be compressed before transport across a bandwidth-constrained wireless channel. Given the emphasis on mobile wireless systems in the past, the media element that has received the most attention in the context of wireless multimedia is speech [6.155]. This is natural because the most widely deployed wireless multimedia system today is cellular telephony, which is a fairly limited bandwidth system. There has also been a great deal of interest in wireless video, given the increased bandwidth capabilities of Universal Mobile Telecommunications Systems (UMTS). The two video compression standards that are most relevant to these systems are MPEG-4 and H.263, both of which have been evaluated for uses in GSM systems. Because of the unreliable nature of wireless networks, it has become important to build source coding schemes that are robust to channel errors. Scalable compression schemes that offer graceful degradation with loss of data have became popular.

Audio and graphics are two source elements that have not received extensive research in the context of wireless systems. There has, however, been some work on handwriting coding [6.159].

Even with scalable and multiple description-based source-coding schemes, there will still be lost data on wireless systems. Error recovery and concealment at the receiver is therefore an important topic and has received some attention again, primarily for video. These error-concealment techniques rely to a large extent on knowing the underlying source compression technique and exploiting some of the tools that are used therein [6.160].

Most of the wireless systems today also support mobility. The velocity associated with mobility has been one of the key parameters that affect system design. For this reason, many of the approaches to solving channel-related problems associated with mobility have been developed for specific classes of mobile systems: pedestrian (velocity of a few meters/sec), vehicular (velocities of about 100 meters/sec) and high speed (velocities of hundreds of kilometers/sec). Mobility also affects routing and addressing, which have received a significant amount of attention.

6.7.1 Wireless Broadband Communication System (WBCS) for Multimedia

Depending on its applications, there are two distinct approaches to the development of WBCS: Wireless LAN (WLAN) and Mobile Broadband System (MBS). Although the core network dilemma is still going strong between IP and ATM for broadband multimedia services, almost all of the WBCS technology demonstrations are based on ATM technology. ATM as a broadband infrastructure has been designed for multimedia communications to accommodate a variety of data rates, QoS requirements and connection and connectionless paradigms. It is quite natural to assume a combination of wireless and ATM-based services at the consumer end of a wired network. In order to deliver multimedia traffic across broadband wireless networks, we need to have sufficient bandwidth and be able to support service-specific QoS requirements concerning delay, delay variation and packet loss on a per-connection basis.

The radio physical layer is essentially the soul of any wireless network. Ideally, one wants to find a radio physical layer technology that is spectrum efficient, minimizes the radio overhead and is robust in both indoor and outdoor environments. Because of various channel impairments, it is very hard to get an optimal radio physical layer.

The wireless broadband air interface will demand a relatively large frequency band to support bit rates in excess of 2 Mb/s. This type of allocation is hard to find lower than 3 GHz, and the availability of bandwidth becomes easier on higher frequencies, but at the cost of complex and expensive techniques [6.156]. Because at higher frequencies the path loss is greater, line of sight operation becomes important and wall penetration becomes a challenge for WLANs.

In the wireless environment, the transmitted radio signal is subject to various time-varying impairments that arise from inherent user mobility and unavoidable changes related to the movement of the surrounding environment. This results in fading and shadowing effects. Another problem is the presence of multipath propagation, leading to fading and time delay spread, which give rise to Intersymbol Interference (ISI) that can strongly increase the BER. One way to overcome fading induced impairments is to use antenna diversity techniques. This is a useful concept for capacity enhancement. The combination of antenna diversity and equalization has the potential to offer significant performance and capacity gains. Another way to improve spectrum efficiency is through power control [6.161].

The wireless transmission medium is a shared radio environment. Therefore, coordinated scheduling of transmission by a control access point can be used to maximize throughput. The major issues are to define a flexible air interface and efficient error control and traffic scheduling algorithms. Data Link Control (DLC) is the core for multiplexing services with varying QoS demands. A generic wireless DLC consists of a flexible packet access interface, delay- oriented and delay-variation-oriented scheduling for multimedia traffic and error control per service requirement.

Multiple Access (MA) protocol is required to minimize or eliminate the chance of collision of different information bursts transmitted from different users. Ideally, the desired MA scheme should be insensitive to the channel impairments. The services provided to the user must satisfy certain quality requirements no matter how bad the channel is. In addition, a good MA protocol can improve system capacity and can lower the system cost. Hence, the MA scheme is a very important design issue for efficient and fair use of the available system resources. In order to provide multimedia services under the limited bandwidth constraint, a sophisticated MA scheme is crucial to cope with various traffic characteristics. Voice is delay sensitive, but relatively loss insensitive; data are loss sensitive, but delay insensitive and voice data rate is generally much higher than either voice or ordinary data and is also delay sensitive. In wireless communication, the multiple access channel can be shared by a large number of users using an MA scheme. There are three types of MA schemes: Frequency Division Multiple Access (FDMA), TDMA and Code Division Multiple Access (CDMA). FDMA assigns a unique frequency band to each user. TDMA assigns access in time slots, and CDMA assigns a unique code using a spread spectrum technique [6.162].

Whereas the DLC layer is used to enhance the transport capability of the physical layer, the LLC layer is used to improve error performance. Error control is typically achieved by coding and/or retransmission. A trade-off between coding and retransmission has to be optimized for the efficient transmission of data across the air interface. For error control, FEC and Automatic Repeat Request (ARQ) are very effective in improving QoS parameters.

Channel allocation is an important issue in radio resource management. It involves allocating the radio resources systemwide to achieve the highest spectrum efficiency. There are two basic forms of channel allocation, namely Fixed Channel Allocation (FCA) and Dynamic Channel Allocation (DCA) [6.154].

6.7.2 Audiovisual Solutions for Wireless Communications

There has been a great deal of standardization and research effort in the area of audiovisual coding for wireless communications. Many papers and reviews have been published covering different aspects of audiovisual coding for wireless communications [6.155, 6.163, 6.164]. In general, previous reviews have focused on wireless speech-coding schemes and error-resilient video standards. Meanwhile, a great deal of research has been conducted in the area of joint source-channel coding for error-prone networks.

In the area of wireless audio, the focus has been on the development of speech-coding solutions for cordless systems, cellular telephony services and emerging personal communication services [6.155]. Because of the time-varying impairments that characterize wireless communication channels, special attention has to be paid when coding any multimedia signal to be delivered across such networks. In particular, compressed image or video signals can experience severe degradation if transmitted across error-prone channels. This is due mainly to the following:

- The use of Variable Length Coding (VLC) in compressed bit stream
- The development of prediction-based coding needed for eliminating both spatial and temporal redundancies in the original signal

Channel errors affecting a VLC could result in a loss of synchronization at the decoder.

Example 6.3 The picture is divided into equal regions of pixels where each region is referred to as a GOB. The example assumes that a synchronization code is used in the bit stream at the beginning of every GOB. Therefore, an error inside the compressed bit stream could damage the remainder of the GOB being decoded until synchronization is achieved at the beginning of the next GOB. An example illustrating the impact of an error in the compressed data on the region of a picture in the pixel domain is shown in Figure 6.56.

Example 6.4 When interpicture prediction-based coding is employed, major degradation in a video sequence can be observed if a reference frame (for example, an intracoded picture) experiences any errors. In this case, when a corrupted picture is used to predict another picture, the error will propagate in the video sequence until the affected area is refreshed by transmitting intracoded blocks or a whole new intracoded picture. An example of an error-propagation scenario in a video sequence is shown in Figure 6.57.

RVLCs enable the receiver to decode the bit stream in a backward manner starting from the next resynchronization marker after an error. Therefore, for any bit stream segment located between two consecutive synchronization markers, RVLCs could assist (Figure 6.58) the decoder in isolating and consequently discarding the region of the bit-stream segment experiencing one or more errors.

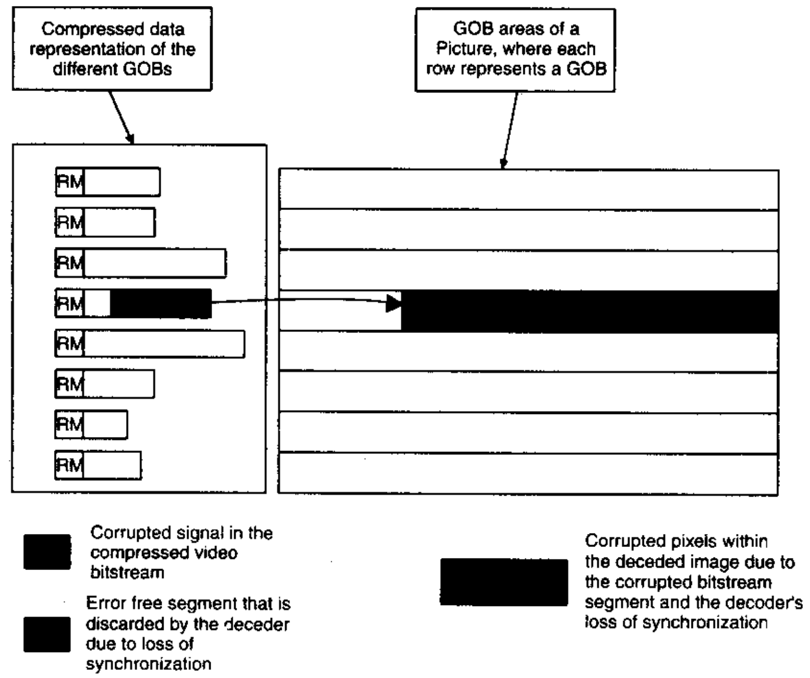

Figure 6.56 The impact of an error in the compressed data on the region of a picture in the pixel domain.

The data segment for the case with RVLC is shown wider (that is, more bits) than the case without RVLCs because some coding efficiency is lost when using RVLCs.

In general, adding any robustness to the compressed signal in the form of resynchronization bits or using RVLCs reduces the coding efficiency. In order to maintain a good balance between coding efficiency and error resilience, other mechanisms have been proposed in conjunction with standard-compliant bit streams. For example in Reyes et al. [6.164], a transcoding scheme is used to increase the robustness of an H.263 signal transmitted across a combined wired-wireless network. The transcoding mechanism, which is employed at the boundary between the wired and wireless segments of the combined network, is designed to improve the resilience of the video signal to errors while minimizing loss of coding efficiency.

In Joint Source Channel (JSC) coding, the time-varying characteristics of an error-prone channel are taken into consideration when designing the source and channel coders of a wireless system [6.165 through 6.172].

Figure 6.59 shows a generic model of a visual source-channel coder. The source signal could be either a still image or a video sequence. For video sequences, the source signal could be either an original picture (for an intracoded frame) or a residual signal representing the difference between an original picture and a prediction of that picture (for example, for motion-compensated prediction). The source signal usually undergoes an orthogonal transform that provides clustering of high-energy coefficients in a compact manner. The DCT is an example of such a

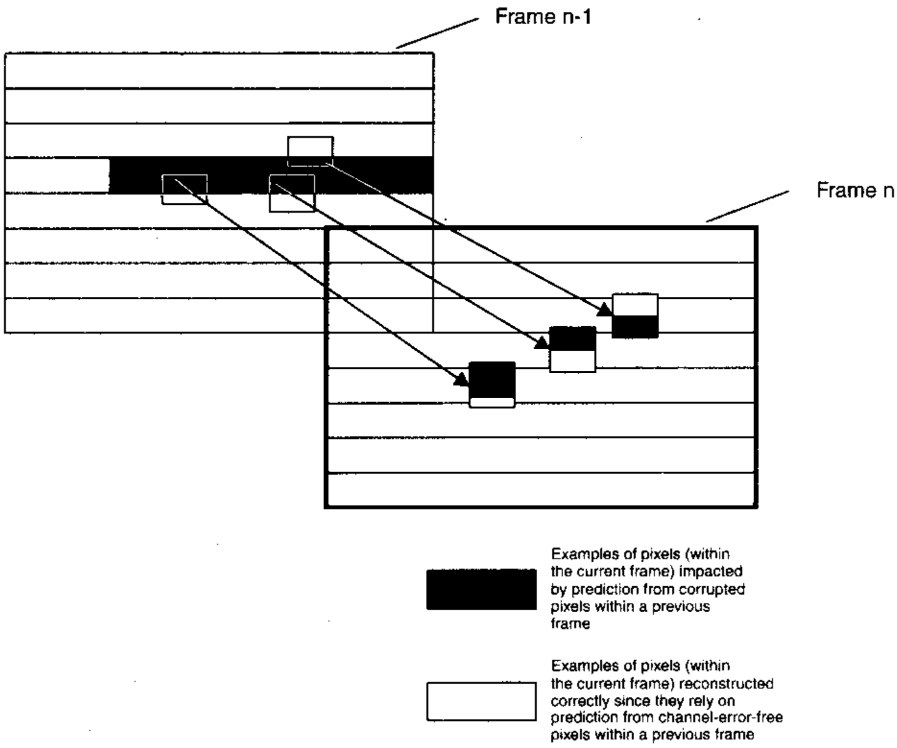

Figure 6.57 Error propagation due to interframe prediction coding.

transform. DCT has been proposed in the context of optimized JSC coding [6.167, 6.173]. Meanwhile, wavelet transform and subband-based video coding have also been proposed for many JSC solutions [6.165, 6.168, 6.169, 6.170]. Moreover, wavelet transform-based image compression has been adopted in MPEG-4 as the basis for a still-image texture-coding tool and also in JPEG 2000. Wireless and mobile applications are among key target application areas for both of these standards.

The second stage of the generic JSC model is the classification and grouping of the transform coefficients. This stage is needed for more efficient and robust quantization and coding (source and/or channel) of the coefficients. In Li and Chen [6.167], a different type of classification and grouping is employed. Although a block-based DCT is used, the DCT coefficients are grouped on the basis of their frequencies into B subband images (subsources). Therefore, for NxN DCT blocks, there are N^2 subsources. Each subsource contains all coefficients with the same frequency (for example, the DC coefficients). This enables the sender to allocate different amounts of bits for the different subsources. Most of wavelet transform-based video-coding solutions proposed for error-prone channels are based on grouping the wavelet coefficients using EZW [6.174, 6.175]. An improved variation of the EZW algorithm is known as Set Partitioning in Hierarchical Trees (SPIHT) [6.175]. In EZW coding, efficiency is achieved on the basis of the

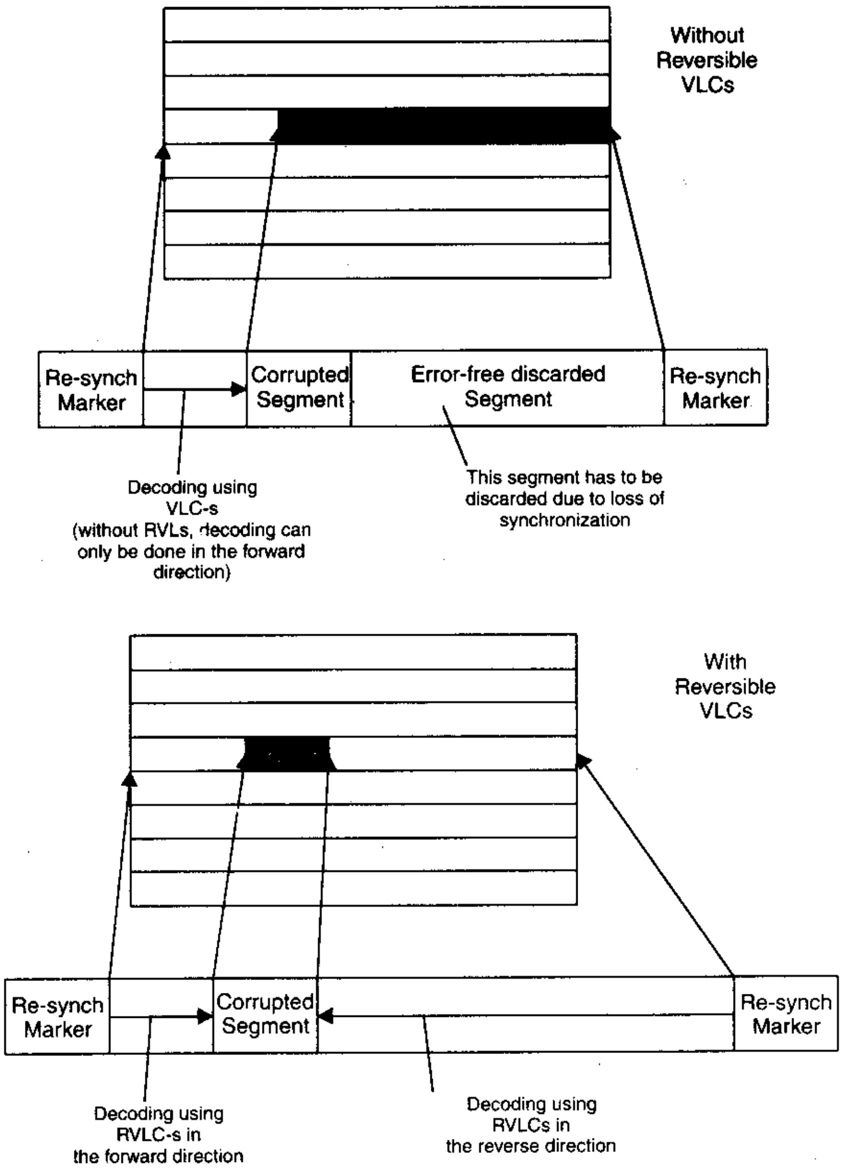

Figure 6.58 Benefits of using RVLCs [6.163]. ©1997 IEEE.

hypothesis of decaying spectrum. The energies of the wavelet coefficients are expected to decay in the direction from the root of a spatial orientation tree toward its descendents. If the wavelet coefficient c_n of a node n is found insignificant (relative to some threshold $T_i = 2^k$), it is highly probable that all descendants D(n) of the node n are also insignificant (relative to the same

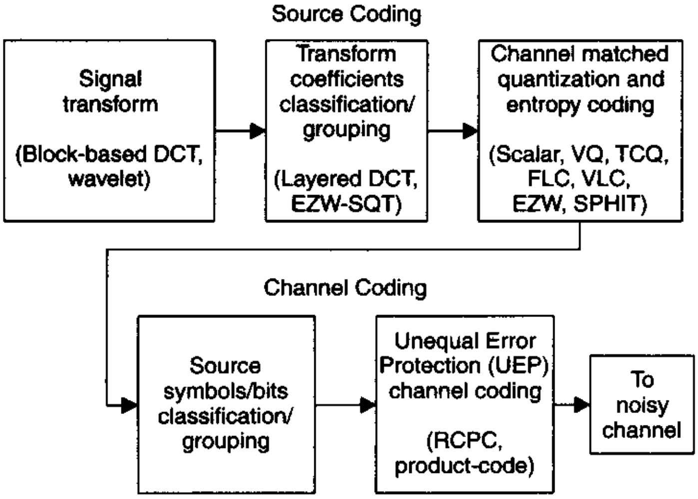

Figure 6.59 Generic model of JSC coding system for a noisy wireless network [6.165]. ©1996 IEEE.

threshold T_k). If the root of a tree and all of its descendents are insignificant, this tree is referred to as a Zero Tree (ZTR). If a node n is insignificant (i.e., $|c_n| < T_k$) but one (or more) of its descendants is (are) significant, then this scenario represents a violation of the decaying spectrum hypothesis. Such a node is referred to as an Isolated Zero Tree (IZT). In the original EZW algorithm, a significant coefficient c_n (i.e., $|c_n| > T_k$) is coded either positive (POS) or negative (NEG) depending on the sign of the coefficient. If $S(n,T_k)$ represents the significance symbol used for coding a node n relative to a threshold $T_k = 2^k$, then

$$S(n,T_k)=\begin{cases} ZTR & if\ |c_n|<T_k\ and\ \max\limits_{m\in D(n)}(|c_n|<T_i) \\ IZT & if\ |c_n|<T_k\ and\ \max\limits_{m\in D(n)}(|c_n|\geq T_k) \\ POS & |c_n|\geq T_k\ and\ c_n>0 \\ NEG & |c_n|\geq T_k\ and\ c_n<0 \end{cases} \quad (6.$$

The coding procedure used for coding the significance of the wavelet coefficients is referred to as significance map coding.

The third stage of the generic JCS model is the quantization and entropy coding of the classified transform coefficients. Fixed-length entropy coding (that is, using Fixed Length Codes [FLCs]), VLCs, hybrid fixed-variable length coding or arithmetic entropy-based coding mechanisms are normally used in conjunction with some type of quantization. In addition to scalar quantization, VQ [6.176] and TCQ [6.177] are among the popular techniques proposed for wireless video.

After the source encoder generates its symbols, the channel coder provides the necessary protection to these symbols prior to their transmission across error-prone networks. One popular approach that has been used extensively is the UEP paradigm. The UEP paradigm enables the channel coder to use different levels of protection depending on the channel condition. In addition, under UEP, the channel coder can provide different levels of protection for the different source symbols depending on their importance. The importance of a source symbol can be measured on the basis of the amount of distortion when that symbol is corrupted by a channel error. For any source-driven UEP approach, there is a need for a classification process for the source symbols. In Li and Chen [6.167], a simple layered approach is used to classify the different bits of the Uniform Threshold TCQ (UTTCQ) coder. The higher the layer in which the bits are located (within the trellis) means the more important they are. All bits in the same layer are treated in the same way (in terms of importance) and therefore are grouped into one data block. Then different levels of protection are used for the different data blocks. Rate-Compatible Punctured Convolutional (RCPC) codes were invented to provide a viable and practical channel-coding solution for the UEP paradigm [6.178, 6.179]. RCPC codes are generated by a single (channel) encoder decoder puncturing the code at the output of a convolutional coder. Because of its flexibility and low complexity, the RCPC channel coder has been extremely popular for wireless video transmission. Consequently, it has been employed for source driven, channel driven or both source and channel UEP.

6.7.3 Mobile Networks

Broadband wireless communications have gained increased interest during the last few years. This has been fueled by a large demand on high-frequency use as well as a large number of users requiring simultaneous high-data-rate access for the applications of wireless mobile Internet and e-commerce. The convergence of wireless mobile and access will be the next storm in wireless communications, which will use a new network architecture to deliver broadband services in a more generic configuration to wireless customers, and it will support value-added services and emerging interactive multimedia communications. Large bandwidth, guaranteed QoS, and ease of deployment coupled with recent advancements in semiconductor technologies make this converged wireless system a very attractive solution for broadband service delivery.

One of the most interesting activities in mobile communications today is the development and standardization of so-called Third Generation (3G) mobile systems known as UMTS in European Telecommunication Standards Institute (ETSI) and International Mobile Telecommunication (IMT) IMT-2000 in ITU. The first system was deployed in 2001. The 3G systems will provide the user with higher data rates than the current second generation systems, such as GSM. Largely because of the higher data rates, the 3G systems are also expected to enable the use of multimedia applications including, for example, video content. IMT-2000 is expected to play the key role of the mobile telecommunication infrastructure for providing multimedia services supported by user bit rates up to 2 Mb/s.

At the same time, the second generation systems are developing fast, and it is now obvious that many of the 3G applications will be realized already in second generation systems, and the transition to 3G systems, although bringing a performance enhancement, will be smooth from the application point of view.

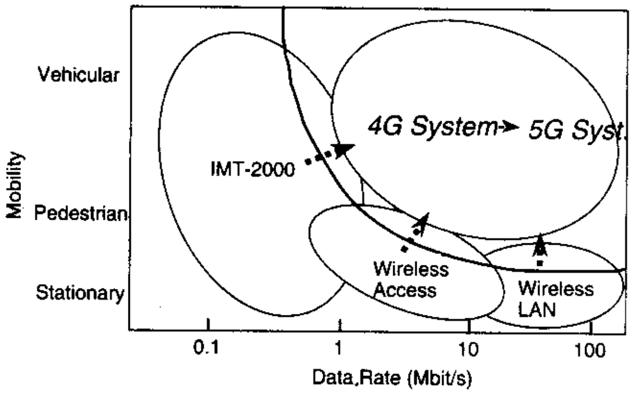

Figure 6.60 Targets of mobile communication systems beyond IMT-2000.

In the history of mobile systems, the 1980s are the analog systems era. Because it was the beginning of mobile communication services, the analog systems are called the first generation systems. Roughly 10 years ago, digital cellular and cordless services were started. The 1990s are called the second generation era for mobile systems. Like in the first generation systems, telephone is the major service in the second generation systems. Multimedia mobile services and worldwide roaming capability are requested for mobile systems. To cope with these demands, IMT-2000 has been standardized as the 3G system, and the services were started in 2001. Figure 6.60 shows the targets of mobile communication systems beyond IMT-2000. IMT-2000 will support the data rates of up to 2 Mb/s. In the high mobility environment, it will be 144 Kb/s. Namely, IMT-2000 offers full mobility capability, but its data rates are limited. As shown, there are different approaches to the fourth generation (4G) system, research on wireless access and WLAN systems. Under the limited mobility condition, the targets of these systems are on even higher data rates corresponding to the broadband services in the fixed networks. The 4G system has to support next generation applications and to be transparent with fixed network. The network of 4G systems may be new one with IP capability. Seamless service area and high-data-rate services should be supported. A flexible mobile terminal with multimode functions will be needed.

Using GSM as an example, at least two clear trends can be seen in the continuous enhancement of the system:

- The basic speech-service is developing with enhancements in quality/capacity performance.
- The available data rates are increasing, making live video transmission and multimedia realistic.

Speech Transmission in GSM

The dominant service in mobile communication systems is the basic telephony service. GSM speech service has been improved in two different standardization items since 1995: the Enhanced Full Rate (EFR) and the Adaptive Multirate (AMR) codecs [6.180].

The GSM EFR codec provides speech quality equivalent to that of a wireline telephony reference (ADPCM 32 Kb/s). The EFR codec uses a 12.2 Kb/s bit rate for speech coding and 10.6 Kb/s for error protection adding up to the GSM full rate channel total of 22.8 Kb/s. Speech coding is based on the Algebraic Code Excited Linear Prediction (ACELP) algorithm. The codec provides substantial quality improvement compared to the existing GSM Full Rate (FR) and Half Rate (HR) codecs. The EFR codec provides wireline quality for the most typical error conditions as well as for background noise and mobile-to-mobile calls. Depending on the implementation platform for a mobile handset, the computational complexity of the EFR codec is within 15 to 20 MIPS. The algorithmic delay of the codec is 20 ms [6.181].

The performance of EFR could be improved by using a different capacity allocation between source and channel coding in severe channel error conditions. The GSM AMR codec would operate in FR (22.8 Kb/s) and HR (11.4 Kb/s) channels using multiple bit rates and adapting the source coding and channel-coding bit-rates according to the estimated quality of the radio channel. By switching the codec to operate in the FR channel during good channel conditions, the AMR codec can also provide channel capacity gain over the EFR codec. Compared to the earlier GSM speech codecs, the ARM codec also requires the specification of link adaptation with inband codec mode control and transmission of channel quality measurement data [6.182]. The channel error performance was studied by using a subjective listening test and error-free 16 Kb/s (G.728) Low Delay CELP (LD-CELP) codec as a high-quality reference codec. The GSM FR codec was also included as an additional reference for a subset of test conditions. The mode adaptation algorithm has also been tested with dynamic channel models where the error conditions change during the test sample. The improvement is substantial in all conditions validating the AMR concept.

Video Across GSM

The GSM system presently offers circuit-switched data rates of 9.6 Kb/s which is usually not adequate for real-time transmission of video. However, the specifications for High-Speed Circuit-Switched Data (HSCSD) [6.183] have already been completed, and deployment of these higher data rates for GSM started in 1999. The data-rate enhancements include an increase of the single slot data rate from 9.6 Kb/s to 14.4 Kb/s, as well as the use of multiple data transmission slots for one connection. For example, one practical configuration will use two data slots providing an aggregate full duplex data rate of 28.8 Kb/s. This rate will already be sufficient for simple video telephony as well as for some asymmetric video applications. Even with increased data rates, the acceptance of low-bit-rate video quality will remain an important issue. It is essential to continue improvement in the video compression efficiency. Providing high data rate services for wide area coverage will always carry a cost, and the capacity pressure will favor efficient compression.

In general, the mobile video coder has the same operating principle as most other compressed video coders as shown in Figure 6.61. The prediction error $e_n(x,y)$ is compressed and sent to the decoder together with motion vectors. To indicate that the compression of the prediction error is typically lossy, the compressed prediction error is denoted as $\tilde{E}_n(x,y)$. In the decoder the nth frame of the sequence is reconstructed by predicting each segment and then by adding the received prediction error, that is,

$$\tilde{I}_n(x,y) = P_n(x,y) + \tilde{E}_n(x,y) \quad (6.19)$$

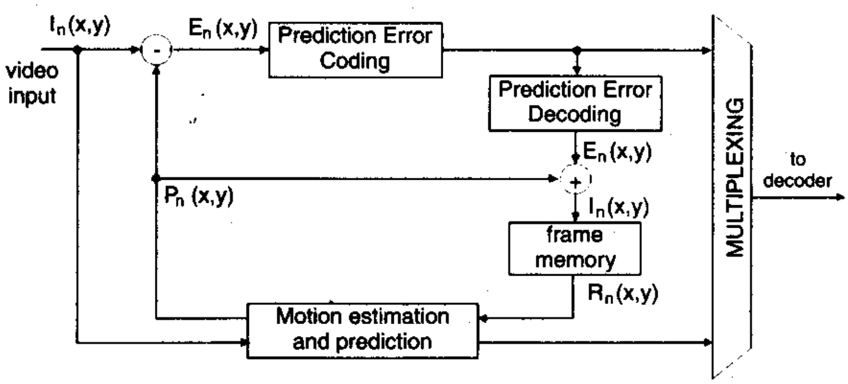

Figure 6.61 Block diagram of the mobile video encoder [6.184]. ©1997 Elsevier.

The block diagram of the motion estimation module is shown in Figure 6.62.

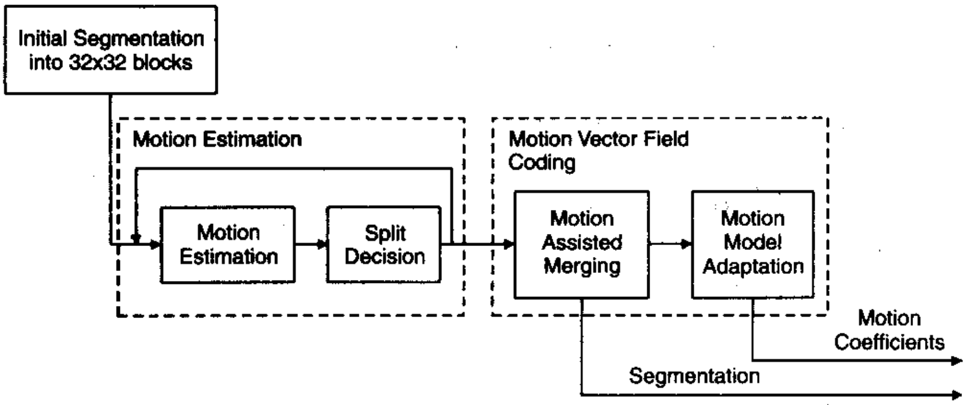

Figure 6.62 Block diagram of the mobile video motion vector field estimation and encoding [6.184]. ©1997 Elsevier.

The segmentation starts by splitting 32x32 blocks in quad-tree fashion as long as sufficient motion estimation performance improvement is obtained. The resulting segments are again merged together based on the motion estimation error. The affine motion model of each segment is then reduced to contain the smallest possible number of motion coefficients without compromising motion estimation performance. The improved compression performance of mobile video is a significant benefit for mobile applications. However, real-time video across cellular phones typically also requires error-resilience enhancement both to the system level as well as to the video codec itself.

Mobile ATM

The item, which defines the design functions of control/signaling, is called mobile ATM. In WATM networks, a mobile end-user establishes a VC to communicate with another user, either a mobile or an ATM end-user. When the mobile end-user moves from one AP to another AP, proper handover is required. To minimize the interruption to cell transport, an efficient switching of the active VC from the old data path to new data path is needed. Also, the switching should be fast enough to make the new VCs available to the mobile users. During the handover, an old path is released, and a new path is then reestablished. In this case, no cell is lost, and cell sequence is preserved. Cell buffering consists of uplink buffering and downlink buffering. If VC is broken when the mobile user is sending cells to APs, unlinking buffering is required. The mobile user will buffer all the outgoing cells. When the connection is up, it sends out all the buffered cells so that no cells are lost unless the buffer overflows. Downlink buffering is performed by APs to preserve the downlink cells for sudden link interruption congestion or retransmissions. It may also occur when the handover is executed [6.185].

When a connection is established between one mobile ATM endpoint and another ATM end point, the mobile ATM endpoint needs to be located. There are two basic location management schemes: the mobile scheme and the location register scheme. In the mobile scheme, when a user moves, the reachability update information only propagates to the nodes in a limited region. When a call is originated by switching in this region, it can use the location information to establish the connection directly. If a switch outside this region originates a call, a connection is established between this switch and the mobile's home agent, which then forwards the calls to the mobile. This scheme decreases the number of signaling messages during a local handover. In a location register scheme, an explicit search is required prior to the establishment of connections. A hierarchy of location registers, which is limited to a certain level, is used [6.186].

Mobile IP

The evolution of mobile networking will differ from that of telephony in some important respects. The end points of a telephone connection are typically human. Computer applications are likely to involve interconnections between machines without human interruption. Obvious examples of this are mobile computing devices on airplanes, ships and automobiles. Mobile networking may well also come to depend on position-finding devices, such as a satellite global positioning system, to work in tandem with wireless access to the Internet. There are still some technical obstacles that must be overcome before mobile networking can become widespread. The most fundamental is the IP, the protocol that connects the networks of today's Internet, and routes packets to their destinations according to IP addresses. These addresses are associated with a fixed network location much as a nonmobile phone number is associated with a physical jack in a wall. When the packet's destination is a mobile node, this means that each new point of attachment made by the node is associated with a new network number and, hence, a new IP address. Mobile IP is a proposed standard protocol that builds on IP by making mobility transparent to applications and higher-level protocols like TCP.

Mobile IP (RFC2002), a standard proposed by a WG within the IETF, was designed to solve this problem by allowing the mobile node to use two IP addresses: a fixed home address and a core address that changes at each new point of attachment [6.187, 6.188]. There is a great deal of interest in mobile computing and apparently in mobile IP as a way to provide for it. Mobile IP is the basis either directly or indirectly of many current research efforts and products. For example, the Cellular Digital Packet Data (CDPD) has created a widely deployed communications infrastructure based on a previous draft specification of the protocol [6.189]. In addition, most major router vendors have developed implementations for mobile IP.

IP routes packets from a source end point to a destination by allowing routers to forward packets from incoming network interfaces to outbound interfaces according to routing tables. The routing tables maintain the next-hop (outbound interface) information for each destination IP address, according to the number of networks to which that IP address is connected. The network number that is derived from the IP address typically carries information with it that specifies the IP node's point of attachment. To maintain existing transport layer connections as the mobile node moves from place to place, it must keep its IP address the same. In TCP, connections are indexed by a quadruplet that contains the IP addresses and port numbers of both connection endpoints. Changing any of these numbers will cause the connection to be disrupted and lost. On the other hand, correct delivery of packets to the mobile node's current point of attachment depends on the network number contained within the mobile node's IP addresses, which change at new points of attachment. To change the routing requires a new IP address associated with the new point of attachment. Mobile IP has been designed to solve this problem by allowing the mobile node to use two IP addresses. In mobile IP, the home address is static and is used, for instance, to identify TCP connections. This takes care of address changes at each new point of attachment and can be thought of as the mobile node's topologically significant address. It indicates the network number and thus identifies the mobile node's point of attachment with respect to the network topology. The home address makes it appear that the mobile node is continually able to receive data on its home network when mobile IP requires the existence of a network node known as the home agent. By home network, we mean the network at which the mobile node seems reachable to the rest of the Internet, by virtue of its assigned IP address. Home agent is a node on the home network that effectively causes the mobile node to be reachable at its home address even when the mobile node is not attached to its home network. Whenever the mobile node is not attached to its home network, the home agent gets all the packets to arrive for the mobile node and arranges to deliver them to the mobile node's current point of attachment.

Whenever the mobile node moves, it registers its new care-of address with its home agent. To get a packet to a mobile node from its home network, the home agent delivers the packet from the home network to the care-of address. Further delivery requires that the packet be modulated so that the care-of address appears at the destination IP address. When the packet arrives at the care-of address, the reserve transformation is applied so that the packet once again appears to have the mobile node's home address as the destination IP address. When the packet arrives at the mobile node, addressed to the home address, it will be processed properly by TCP.

Mobile IP is best understood as the cooperation of three separable mechanisms:

- Discovering the care-of address
- Registering the care-of address
- Timing to the care-of address

Mobile IP discovery does not modify the original fields of existing router advertisements, but simply extends them to associated mobility functions. When the router advertisements also contain the needed care-of address, they are known as agent advertisements, which are the procedures by which a mobile agent becomes known to the mobile node. Home agents and foreign agents typically broadcast agent advertisements at regular intervals, for example, once a second or once every few seconds. An agent advertisement performs the following functions:

- Allows for the data detection of mobile agents
- Lists one or more available care-of addresses
- Informs the mobile node about special features provided by foreign agents, for example, alternative encapsulation techniques
- Lets mobile nodes determine the network number and status of their link to the Internet
- Lets the mobile node know whether the agent is a home agent, a foreign agent or both, and therefore whether it is on its home network or a foreign network

After a mobile node has a care-of address, its home agent must find out about it. Figure 6.63 shows the registration process defined by mobile IP for this purpose. The process begins when the mobile node, possibly with the assistance of a foreign agent, sends a registration request with the care-of address information. When the home agent receives this request, it adds the necessary information to its routing table, approves the request and sends a registration reply back to the mobile node.

In mobile IP, foreign agents are mostly passive, relaying registration requests and replies back and forth between the home agent and the mobile node. The foreign agent also decapsulates traffic from the home agent and forwards it to the mobile node [6.191].

Figure 6.64 shows the tunneling operations in mobile IP. The default encapsulation mechanism that must be supported by all mobile agents using Mobile IP is IP-within-IP [6.192]. By encapsulation, we mean the process of incorporating an original IP packet inside another IP packet, making the fields within the original IP header temporarily lose their effects. Using IP-within-IP, the home agent, or the tunnel's service, inserts a new IP header, or tunnel header, in front of the IP header of any datagram addressed to the mobile node's home address. The new tunnel header uses the mobile node's care-of address as the destination IP address, or tunnel destination. The tunnel source IP address is the home agent, and the tunnel header uses 4 as the

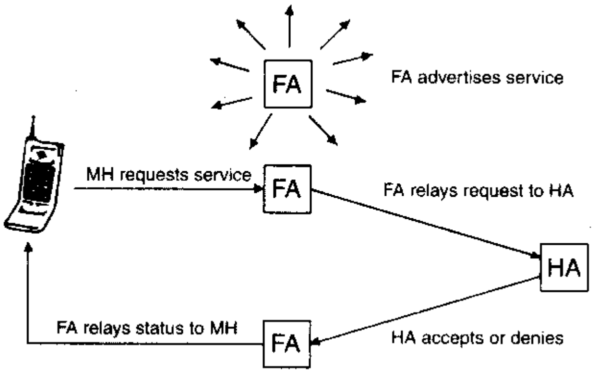

Figure 6.63 Registration operations in mobile IP [6.190]. ©1998 IEEE.

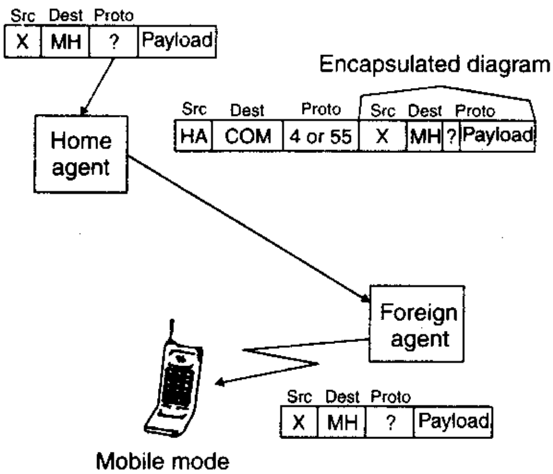

Figure 6.64 Tunneling operations in mobile IP [6.190]. ©1998 IEEE.

higher-level protocol number, indicating that the next protocol header is again an IP header. In IP-within-IP the entire original IP header is preserved as the first part of the payload of the tunnel header. Sometimes the tunnel header uses protocol number 55 as the inner header. This happens when the home agent uses minimal encapsulation instead of IP-within-IP [6.192]. Processing for the minimal encapsulation header is slightly more complicated than that for IP-within-IP because some of the information from the tunnel header is combined with the information in the inner minimal encapsulation header to reconstitute the original IP header. On the other hand, header overhead is reduced. A complete description of the Mobile IP architecture can be found in RFC2002 [6.188]. Related specifications are available in RFCs 2003 through 2006 [6.193]. According to this, mobile node should register with its home agent each time that it changes its care-of address. If the distance between the visited network and the home network of the mobile node is large, the signaling for these registrations may be long. Mobile IP regional registration is shown in Figure 6.65.

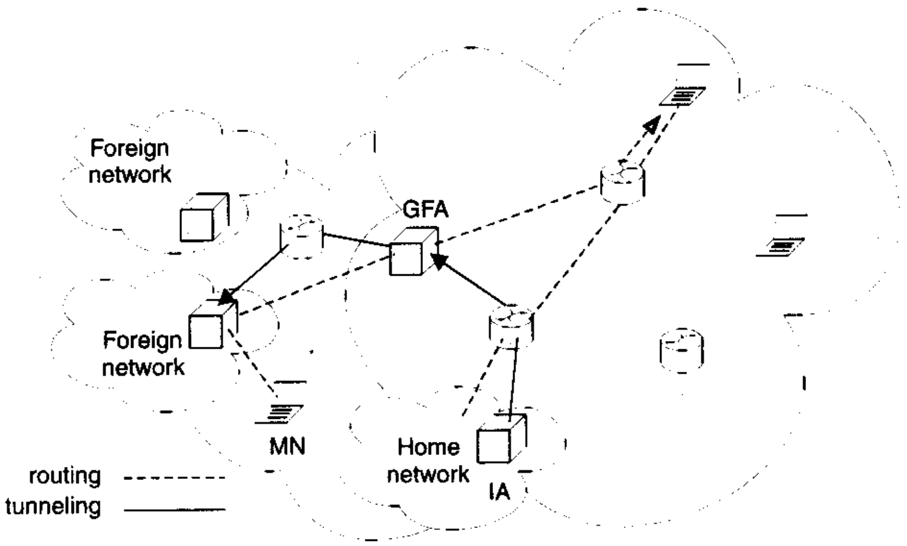

Figure 6.65 Mobile IP regional registration [6.193]. ©2001 IEEE.

With regional registration, when a mobile network arrives for the first time in a new domain, it registers with its Home Agent by means of a local Foreign Agent. However, the Foreign Agent does not register directly with the Home Agent but registers the Mobile Network on a Gateway Foreign Agent (GFA). It is the GFA that registers the Mobile Network with the Home Agent. When the Mobile Network moves from a Foreign Agent to another located within the same domain (behind the same GFA), the new FA simply registers the new care-of address on the GFA. When the regional registration is used, the Corresponding Node (CN) packets are first tunneled from the Home Agent to the GFA and from the GFA to the Foreign Agent. If the reverse tunneling is used, the Foreign Agent forwards Mobile Network datagrams by tunneling to the GFA, and the GFA forwards them by tunneling to the Home Agent that finally forwards them to the CN.

Although the mobile IP architecture has been developed for both IPv4 and IPv6 protocols, some problems arise when mobile IP is used in conjunction with IPv4 private access networks. In the current IPv4-based Internet, a lot of access networks use private address schemes to interconnect the network terminals locally and to connect them to the public Internet. In order to route the packets coming from the private hosts to the public Internet, some mapping/translating mechanisms must be used in order to identify the packets within the public Internet. These schemes are generally called a Network Address Translation (NAT) or in some cases, Network Address and Port Translation (NAPT) [6.194]. There are different reasons for using a private address scheme and NAT mechanisms within a private access network, including the following:

- As a solution to the problem of address depletion in IPv4.

- For security. NAT helps ensure security because each outgoing or incoming request must go through a translation process that also offers the opportunity to authenticate the request or to match it to a previous request.
- For flexibility. NAT conserves the IP addresses within a private network when the network changes its point of attachment to the global Internet. No host reconfiguration is needed.

NAT is included as part of routers and is often part of a corporate firewall. Although NAT has many advantages, it has some problems when it is used in conjunction with nonfriendly protocols and/or applications [6.195]. Particularly, the MIP architecture does not work correctly in a NAT-based access scenario.

Wireless Multimedia Delivery

In the year 2001, the Wireless Multimedia Forum Technical Working Group (WMF TWG) [6.196] published its first document, which recommends technologies, formats and protocols that can be used by the various supply-chain members in the streaming multimedia wireless space. Equipment makers, content developers and service providers that build products conforming to the specifications in Recommended Technical Framework Document (RTFD) Version 1.0 will enable equipment from many vendors to interoperate and will allow software interfaces to be interchangeable across networks [6.197].

Achieving such consensus will accelerate the market for multimedia content and services. For example, having a common technology framework for wireless multimedia delivery will reduce the number of multimedia platforms that content providers will have to support. This should hasten their time to market with new content and services and step up the pace at which their content can reach a broad, far-flung user audience.

The RTFD Version 1.0 standards specification document defines the compression, session initiation/call setup, file format and streaming mechanisms to be used between content-creation subsystems, multimedia distribution servers and wireless multimedia terminals in a streaming multimedia network system. RTFD implementations have resulted in the delivery of interoperable streaming multimedia services to any mobile device in 2001.

New mobile services could include the delivery of news, weather, stock and sports updates to mobile users. In addition, traveling parents could receive clips of a child's soccer game or performance in the school play. Geographic location services could be combined with dating services whereby handheld users could receive a multimedia profile of a dating service candidate who lives in the geographic ballpark of the user's location. Children in Japan are already using cell phones to send animated multimedia greetings to one another, and interactive games that could be streamed among participating users across wireless networks are in development in companies across the globe.

In recommending specifications for common use in wireless multimedia networks, the WMF cooperates with related worldwide standards bodies such as the Third Generation Partner-

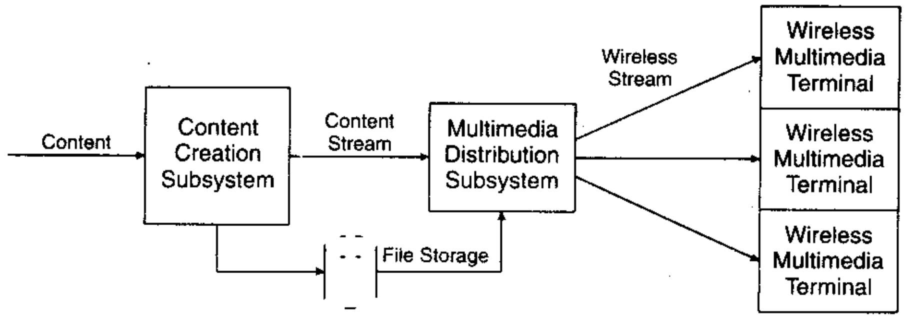

Figure 6.66 Wireless multimedia network system.

ship Project (3GPP) and the ITU. RTFD Version 2.0, already in development, will recommend standards for additional mobile streaming media capabilities, such as QoS, scene description (where graphics and text should appear, relative to multimedia content) and billing capabilities. It will also recommend standards for delivering downloadable multimedia content, such as video email.

RTFD Version 1.0 explicitly addresses the Streaming Multimedia (SMM) application, which includes both on-demand and live streaming using voice and video as the primary media types. The components of an SMM system include the following:

- Content-creation subsystem
- Multimedia distribution
- Wireless multimedia terminals

The content-creation subsystem is responsible for converting raw or compressed media content stored in a file or captured in real time to a content stream suitable for delivery. It then forwards it on to the multimedia distribution server. To do its job, the content-creation subsystem makes use of certain compression technologies and must format the files. A wireless multimedia network system is shown in Figure 6.66.

Content is generated by the content-creation subsystem, distributed to users by the multimedia distribution servers and displayed by wireless multimedia terminals.

The job of the multimedia distribution server, after it has received the multimedia content from the content creation system, is to stream live or stored content to wireless multimedia terminals. RTFD Version 1.0 defines streaming communication between the content-creation subsystem and the multimedia distribution server only for stored content. This multimedia distribution server also can manipulate or repurpose content.

The role of the wireless multimedia terminal is to receive streamed multimedia content from the multimedia distribution server and display it to the user. For streaming media applications, the content may be either live or on demand.

SIP in Mobile Environment

The most important SIP operation is that of inviting new participants to a call. A user first obtains an address where the user is called to translate this address into an IP address where a server may be found. After the server's address is found, the client can send an initiation message to the server. However, as the server that receives the message is not likely to be the host where the user to be invited is actually located, we need to distinguish between different server types that a complete SIP implementation should fulfill. A proxy server receives a request and then forwards it toward the current location of the caller, either directly to the caller or to another server that might be better informed about the actual location of the caller. A redirect server receives requests and informs the caller of the next-hop server. The caller then contacts the next-hop server directly. A user agent server resides on the host where the caller is actually located. It is capable of querying the user about what to do with the incoming call, that is, accept, reject or forward and send the response back to the caller. To assist the end systems in locating their requested communication partner, SIP supports a further server type called register server. It is mainly thought to be a database containing locations as well as user preferences as indicated by the user agents [6.197].

Mobile IP is a widely propagated protocol for supporting mobile communication. With MIP, an end system can be reached in different networks under the same address. However, MIP was primarily designed for TCP communication because it allows for communicating TCP end systems to maintain a connection. This connection is maintained even if there is only one directing the traffic from the calling node to a home agent that maintains location information about the mobile node.

The home agent accepts the TCP connection on behalf of the mobile node and tunnels the traffic either directly toward the mobile node or to a foreign agent that forwards the traffic to the mobile node. The mobile node can communicate directly with the calling node. This kind of communication causes a so-called *triangular routing*, which increases the total end-to-end delay. Further, tunneling the traffic between the home agent and the mobile node increases the amount of consumed bandwidth. This increased complexity is necessary in order to avoid the need for changing the TCP stack at the end systems so as to accommodate for the case of a change in the address of the communicating end systems. A further drawback of MIP is that it uses network terminal addresses to identify the end systems. Although this is appropriate for achieving terminal mobility, it does not easily allow for personal mobility.

SIP provides for personal mobility by using addresses similar to the email addresses that identify a person and not a device. To map this address to a network address, SIP uses the REGISTER method. The same method can be used for providing mobility as well. This means that, whenever the user changes location or device, he can register with the new address. Additionally, the mobile end system informs the other side about the change in the network address. Multimedia communication initiated through SIP is in general based on UDP. Thus, a change in the network address to which to send data does not cause problems with the state of management at the UDP sender as is the case for TCP senders.

Both SIP and Mobile IP share the same problems of authentication and authorization of mobile users. A user roaming into a foreign network needs to authenticate itself to receive access to the network. A harmonization between SIP and Mobile IP might become a promising solution for providing secure and authenticated mobility for both UDP and TCP.

Multicast Routing in Cellular Networks

The most significant trends in today's telecommunications industry are the growth of the cellular network and the rapid rise of the Internet. One of the fastest growing sectors in the telecommunications industry is the cellular service, introducing new demands for user and handset mobility in the future network. A similar explosive increase is observed in the number of Internet subscribers. The current Internet architecture offers a simple point-to-point best-effort service. On the other hand, recently several new classes of distributed applications have been developed, such as remote video, multimedia conferencing, data fusion, visualization, and virtual reality. These applications are not only point-to-point with a single sender and a single receiver of data, but also can often be multipoint-to-multipoint with several senders and several receivers of data. There is a widespread agreement that any new network architecture must be capable of accommodating multicast and a variety of QoS. Multicast enables sources to send a single copy of a message to multiple receivers who explicitly want to receive the information. IP multicasting is a receiver-based concept. Receivers join a particular multicast session group [6.198]. The sender does not need to maintain a list of receivers. The effect of mobility specifically on multicast routing can be summarized as the following:

- When the source of a multicast datagram is a mobile host, a copy of the datagram may not reach all multicast group members, making source-oriented protocol efficient.
- Multicast group members move, requiring an easily reconfigurable multicast tree topology.
- Transient loops may form during tree reconfiguration.
- Channel overhead caused by tree reconfiguration updates increases with mobility, network size and membership size.
- Multicasting necessitates sending copies of a message to multiple locations within the static network. To ensure an acceptable reliable delivery, messages may be buffered at multiple Base Transceiver Stations (BTSs).
- A mobile host may experience a delay in receiving a multicast datagram when it enters a cell that has no other group member located in the same cell.

Core-Based Tree (CBT) is designed to construct and maintain a shared-tree architecture that offers improvements in scalability over source tree architectures by a factor of the number of active sources. Shared trees save bandwidth and state compared with source tree [6.199].

CBT is a backbone within connected group nodes called cores. The backbone is formed by selecting one router, called the primary core, to serve as a connection point for the other cores, called secondary cores. A router willing to participate in the multicast session sends a join request toward the closest core. When the join request reaches a core or an on-tree node, a join acknowledgement is sent back along the reverse path, forming a new branch from the tree to the requesting router. If the core that is reached is a secondary core and is off-tree, it connects to a primary core using the same process. In Brown and Singh [6.200], a new protocol for providing reliable multicast message delivery in mobile networks where the mobile multicast groups experience fragment adds and drops is introduced. Also, a hierarchical network structure, where at the lowest level are the mobile hosts roaming between cells, is proposed. At the next level are the Mobile Support Stations (MSSs), one to a cell, which provide Mobile Hosts with connectivity to the underlying network and with one another. At the top level, groups of MSSs are controlled by a supervisor called the supervisor host (SH). The SH is part of the wired network, maintains connections for MHs and is responsible for maintaining the negotiated QoS. SHs connect mobile networks to the fixed networks and communicate among themselves across the fixed networks.

Broadband Wireless Mobile

Convergence of broadband wireless mobile and access will be the next storm in wireless communications. Fueled by many emerging technologies, including digital signal processing, software-definable radio, intelligent antennas, semiconductor devices and digital transceivers, the future wireless system will be much more compact, with limited hardware and more flexible and intelligent software elements. The compact hardware and the very small portion of software will go the way that the computer industry did in the past. A compact multidimensional broadband wireless model will be adopted for system design and implementation.

Wireless mobile Internet will be the key application of the converged broadband wireless system. The terminal will be compatible with mobile and access services, including wireless multicasting as well as wireless trunking. This new wireless terminal will have the following features:

- At least 90% of traffic will be data.
- The security function will be enhanced.
- A voice recognition function will be enhanced.
- The terminal will support single and multiple users with various service options.
- The terminal will be fully adaptive and software reconfigurable.

As wireless communications evolve to this convergence, 4G mobile wireless communications (4G mobile) will be an ideal mode to support high-data-rate connections from 2 to 20 Mb/s based on the new spectrum requirement for IMT-2000 as well as the coexistence of the current spectrum for broadband wireless access. This 4G mobile system's vision aims at the following:

- Providing a technological response to accelerated growth in demand for broadband wireless connectivity
- Ensuring seamless services provisioning across a multitude of wireless systems and networks, from private to public and from indoor to wide area

- Providing optimum delivery of the user's wanted service through the most appropriate network available
- Coping with the expected growth in Internet-based communications
- Opening new spectrum frontiers

The future wireless network should be an open platform supporting multicarrier, multi-bandwidth and multitrend air interfaces, with content-oriented Bandwidth-on-Demand (BoD) services dominant throughout the whole network. In this way, packetized transmission will go all the way from one wireless end terminal directly to another. Figure 6.67 shows a network reference model architecture. The major benefits of this architecture are that the network design is simplified and that the system cost is greatly reduced. The BTS is now a smart, open platform with a basic broadband hardware pipe embedded with a Common Air Interface Basic Input-output System (CAI BIOS). Most functional modules of the system are software definable and reconfigurable. The packet switching is distributed in the broadband packet backbone, or core network, called packet-division multiples (PDM). The wireless call processing, as well as other console processing, is handled in this network. The gateway acts as proxy for the core network and deals with any issues for the BTS, and the BTS is an open platform supporting various standards, optimized for full harmonization and convergence. The terminal mobile station can be single or multiuser oriented, supporting converged wireless applications [6.201].

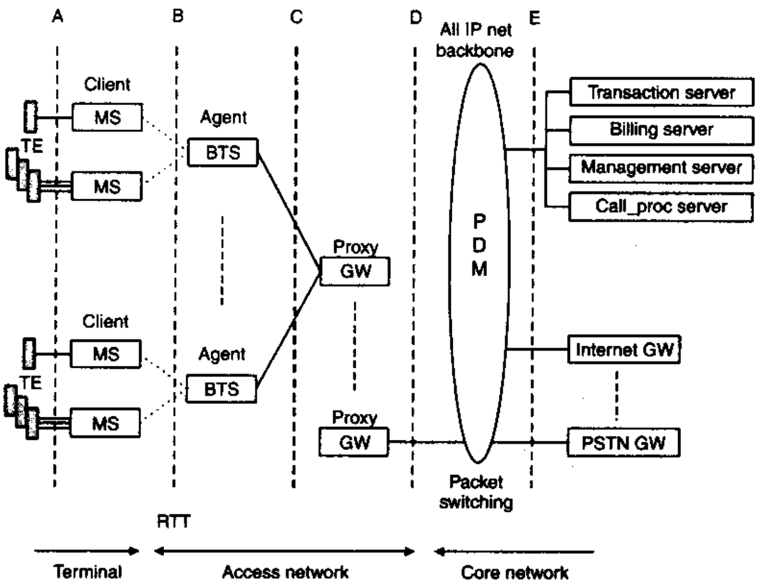

Figure 6.67 A network reference model architecture [6.201]. ©2000 IEEE.

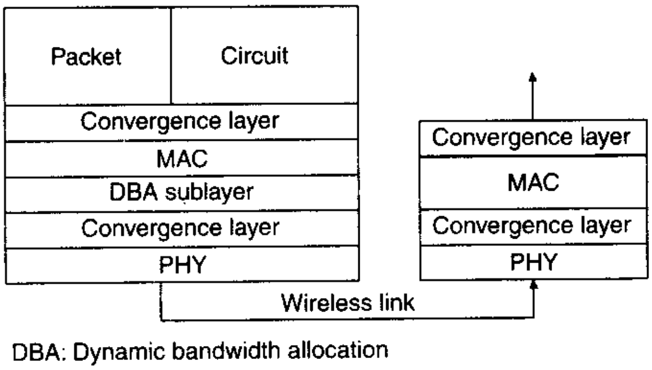

Figure 6.68 General protocol stack [6.201]. ©2000 IEEE.

Considering the signaling protocol, the client/server model is established between a wireless terminal and the core network. The BTS becomes the agent in both directions. Figure 6.68 shows the system protocol stack.

Different services—ATM, IP Synchronous Transfer Mode (STM), and MPEG—can be supported through a service convergence layer. To guarantee wireless QoS and high spectrum use, Dynamic Bandwidth Allocation (DBA) is required through the MAC. The DBA sublayer improves the conventional layer architecture.

The DBA scheduler is the core of the MAC. To realize dynamic resource allocation, this scheduler is essential for the broadband wireless link, which in general helps the following:

- Support class of service offerings
- Provide diagnostic support for all network protocols
- Eliminate the need for traffic shaping
- Increase spectrum use

The transmission convergence layer handles various transmission modulations, error corrections, segmentations and interface mapping of wireless mobile and access in the physical layer.

6.7.4 Broadcasting Networks

A broadcasting system, combined with an interactive channel from a telecommunication system, proves the cheapest and most efficient solution for services that many users share. This holds true as long as the individual information sent to a user stays small compared to the total amount of distributed information. It is more efficient to let the broadcasting system provide a group of users with high capacity rather than divide the capacity into several smaller, individual channels for the same group. Today's market demands global access by multiple users and portable or mobile reception. The cellular phone industry exemplifies this demand. Portable and mobile reception can be offered through a digital terrestrial radio system, or sometimes through satellite. Introducing multimedia in a broadcast system is not straightforward. Existing multimedia systems build on two-way communication links and error-free transmission. However, broadcasters cannot guarantee an error-free or noninterrupted channel. Systems used solely for broadcasting audio or TV were designed with these difficulties in mind. Adding services with multimedia in broadcasting networks demands new concepts that take radio channel characteristics into account.

Figure 6.69 shows an outline of a multimedia system model using broadcast. This model contains three main parts: the content provider, service provider and network provider. Each provider type contains different servers and protocols that include functions, such as an interactive channel, conditional access and system management.

The content provider supplies the information in a multimedia system. The information can consist of audio, video, text, graphics, still pictures or a combination of these. Content providers require appropriate handling of the supplied information in the system and payment for the actual use of the information. Thus, the information can be divided into three different classes:

- Information that the terminal cannot store
- Information that the terminal can store, but not copy
- Information that the terminal can store and copy

A service can consist of a mixture of these information classes. The content provider can offer the same information to a service provider at varying fees, depending on how the service provider uses the information. Content and service providers should agree on the permitted use

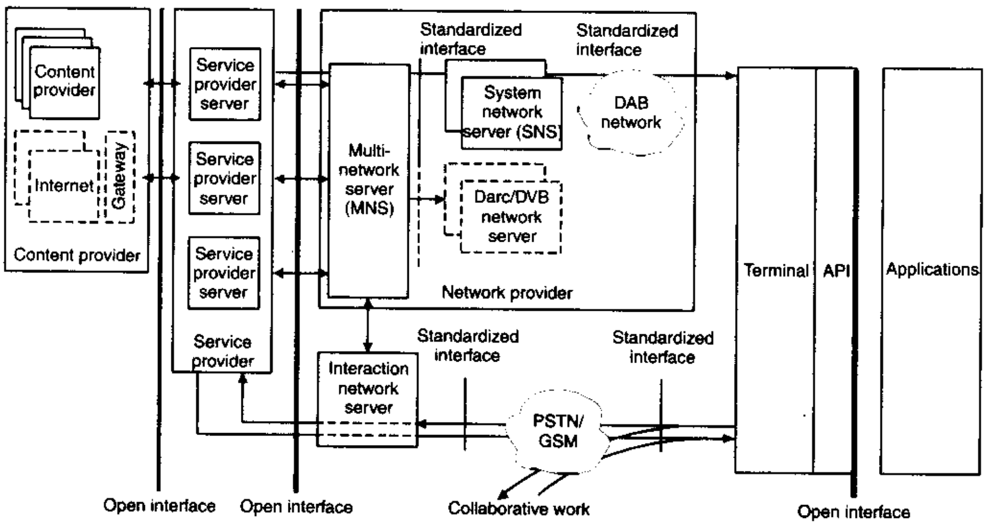

Figure 6.69 Multimedia system model using digital broadcast [6.202]. ©1997 IEEE.

of the information, and service providers can inform end-users of these restrictions. Content providers want to ensure that no one violates the agreed use of the information. Therefore, choosing a system strongly depends on the possibility to protect the content provider's interests.

Service providers create the actual service. They may store the information from the content provider in a server or database and package it into a service (non-real-time services), or they may link the flow of information as a service in itself and distribute it directly without storing it (real-time services). Several layers of service providers could exist, because a service at the terminal end can combine several actual services from different service providers. Service providers must also support subscriber management, service management, information protection mechanisms and billing functions. Service providers receive information that may act as simple channels that transmit whatever content providers supply. In other scenarios, service providers order specific information from content providers, who store the information. If the service is interactive, the service provider may want to access information in distributed databases from various content providers. In other cases, service providers may need to act as buffers to transmit the information, allowing for retransmission. Service providers must also securely handle the content they receive and preserve the content provider's copyright throughout the system.

Selecting which system to use depends on the desired service's capacity requirements. Because network providers can supply several broadcast networks of different types or at different locations, the transport system requires a multinetwork server or router to provide flexible solutions. The multinetwork server can supply system management, select different networks and support other network-independent functions.

Network providers will also have to provide a system network server, which is a unit that translates the information to distribute into the correct format for the transport system. The system network server optimizes the setting of the transport system's available system parameters (such as error protection, repetition and segmentation) in order to adapt to the available transmission channel. It should also support other network-dependent functions.

Network providers must adapt the information transported to several protocol layers according to the system specification. Figure 6.70 illustrates the components of a DAB system's protocol stack. A stream multiplex and fast information channel build the DAB stream. The fast information channel handles multiplex configuration information, which contains information on building the complete multiplex. This includes information such as the number of available audio or data channels, the labels identifying the channels, and descriptions of whether certain channels should link together in the receiver to create a full service. The fast information channel also carries service information describing each service. Streamed audio with program-associated data, a packet mode and proprietary services sit on top of the stream multiplex, possibly encrypted. DAB was developed within the EUREKA147 project and is standardized within the ITU, ETSI and the European Broadcasting Union (EBU). The Coded Orthogonal Frequency Division Multiplex (COFDM) provides the robust transmission channel for mobile reception. The radio frequency signal is wideband (a 1.5 MHz frequency block) with a maximum net bit rate of approximately 1.8 Mb/s, depending on the level of protection.

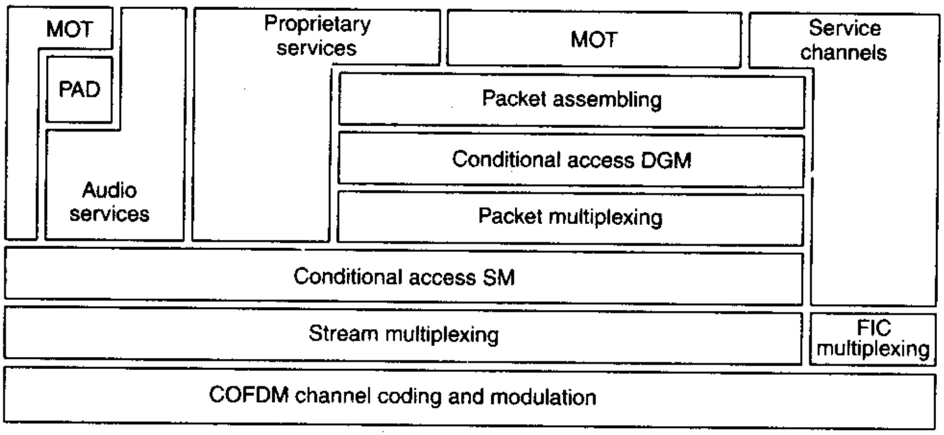

Figure 6.70 Digital audio broadcasting system protocol stack [6.202]. ©1997 IEEE.

DAB has a flexible general-purpose system that reconfigures at any time and supports a wide range of sources, channel-coding options and data services [6.203]. It also incorporates conditional access, such as encryption and addressing, enabling transmission to secluded groups. In conjunction with the DAB specification, a Multimedia Object Transfer (MOT) protocol was developed to support transmitting multimedia objects in the DAB datastream. MOT supports transporting objects and files, segmenting the objects, interleaving on different levels and linking objects in different datastreams. The protocol lets the terminal identify different types of objects, such as JPEG, MPEG or ASCII text so that it can determine whether it has the system resources to handle the object. MOT also includes optional parameters to support applications, such as triggering applications, giving objects time stamps, creating a file name or providing an alternative display mode if a certain decoder is not available.

Network providers require, and sometimes provide, several network-independent functions. These functions should be gathered in an overall logical unit, called a multinetwork server, which supports all the networks. Network providers that manage several networks need to route the information for transmission to the different networks from one or a few central access points. The multinetwork server should support a unified approach so that network providers can bill the service providers and supply the information needed for the service providers to bill the end-users. Scheduling via data carousels, supported by the multinetwork server, can avoid overloading the transmission channels or creating cyclic services. If this functionality accompanies a certain network, it should be distributed to the service network server or a similar unit. Other functions that network providers may support include synchronizing the different data, audio, or video systems before transmission in a single stream in one channel; transmitting the items separately with time stamps (maybe in different channels) and informing the receiver how to assemble all of it. If the latter method is used, the functionality of the transmission system may be chosen to fit each data type work efficiently, but the requirements on the receiver will increase.

In a broadcasting system, service providers have two types of customers: the end-users of the service and the content providers. Service providers should make life as simple as possible for content providers. For example, they should support the content provider's copyright. However, service providers must also try to simplify tasks for end-users. Therefore, determining where to assemble the information becomes an important issue.

A terminal can have many shapes. It can be a receiver dedicated to a particular service, a piece of extraction equipment applied to an ordinary DAB receiver that extracts the service or a DAB receiver (terminal) containing the necessary tools for providing a built-in service. The terminal's basic services include interoperating and presenting data transmitted across air. A large part of the data is presented as soon as it reaches the terminal, like ordinary radio and TV programs. Other parts of the bit stream are immediately disregarded, like the part that represents radio programs other than the ones listened to or data services that the terminal cannot process. Part of the bit stream can be saved onto the hard disk for latter use, either by a broadcast program or interactively by a user.

Digital Video Broadcasting (DVB)

DVB is a transmission scheme based on MPEG-2 video compression utilizing the standard MPEG-2 transmission scheme. DVB provides superior picture quality with the opportunity to view pictures in standard format or wide screen (16:9) format, along with mono, stereo or surround sound. It also allows a range of new features and services including subtitling, multiple audio tracks, interactive content, and multimedia content where, for instance, programs may be *linked* to Web material.

Satellite transmission has led the way in delivering digital TV to viewers. A typical satellite channel has 36 MHz bandwidth, which may support transmission at up to 35 to 40 Mbps using Quadrature Phase-Shift Keying (QPSK) modulation. The audio-video, control data and user data are all formed into fixed-size MPEG-2 transport packets. The complete coding process may be summarized by the following:

- Inverting every eighth synchronization byte
- Scrambling the contents of each packet
- RS coding at 8% overhead
- Interleaved convolutional coding (the level of coding ranges from 1/2 to 7/8, depending on the intended application)
- Modulation using QPSK of the resulting bit stream

The question often arises as to why DVB chose to use MPEG-2. The MPEG-2 coding and compression system was chosen after an analysis and comparison of potential alternatives. Unlike other compression tools that claim to provide greater degrees of compression for given quality, but which are as yet unproven for a wide range of program material or across different broadcasting systems, MPEG-2 is tried and tested. It has been repeatedly shown to be capable of providing excellent quality pictures at bit rates that are practical for the services that will be required. From a commercial point of view, the adaptation of MPEG-2, an existing proven standard, was advantageous because it allowed DVB to concentrate its effort on finding ways of carrying the already well-specified MPEG-2 data packets through a range of different transmission media, including satellite, cable, terrestrial, and so forth. DVB can effectively be regarded as the bridge between broadcasters and the networks across which MPEG-2 data packets can be carried. Another important reason for choosing MPEG-2 was that it includes techniques for the inclusion of Program Specification Information (PSI) for the configuration of decoders. DVB extended these techniques to provide a complete Service Information (SI) capability, enabling receivers to tune automatically to particular services and to decode a mixture of services and service components, including television, sound and data. SI also allows services to be grouped into categories with relevant schedule information, making it possible to provide user-friendly program guides. Another important consideration was that the design of the complete MPEG-2 system makes it possible to freeze the design of a decoder while still retaining the flexibility to make quality improvements at the encoding end of the chain.

DVB has cooperated closely with the Digital Audio Visual Council (DAVIC), whose brief tenure includes the whole range of multimedia transmissions, and many DVB systems have been accepted within the DAVIC standard. MPEG-2 is a video, audio and data coding scheme that can be applied to a range of applications beyond broadcasting, and many multimedia features may, in time, be available from DVB services. DVB systems can naturally carry any or all the items used for multimedia presentations, including text, still pictures, graphics and different types of moving images, and can allow for multimedia extensions to be added. Therefore, DVB members have been focusing on the broadcast market for the immediate commercial future.

The first reason for a broadcaster to select DVB is that DVB makes much better use of available bandwidth. For satellite DTH broadcasters, this advantage is clear. Where a satellite transponder used to carry one analog channel, DVB can offer up to 18 digital channels. DVB-Terrestrial (DVB-T) offers a clearer picture to the end-user, in addition to the capacity for more channels. DVB Cable (DVB-C) offers broadband two-way interactivity. Today, the production, contribution and distribution of television are almost entirely digital. The last step, transmission to the end-user, is still analog. DVB brings this last step into the digital age.

The DVB system has the capability to use a return path between the set-top decoder and the broadcaster. This can be used by a subscriber management system. It requires a modem and the telephone network or a cable TV return path or even a small satellite uplink. This return path can be used for audience participation, such as voting, game playing, teleshopping, telebanking and delivering messages to the decoder. DVB already offers a kind of interactivity without the need for a return path, simply by the breadth of program choices available, for example, multiple sports events and near video on demand.

Data Transmission Using MPEG-2 and DVB

The growing use of multimedia-capable PCs to access the Internet and in particular, the use of the WWW, has resulted in a growing demand for Internet bandwidth. The emphasis has moved from basic Internet access to the exception that connectivity may be provided regardless of the

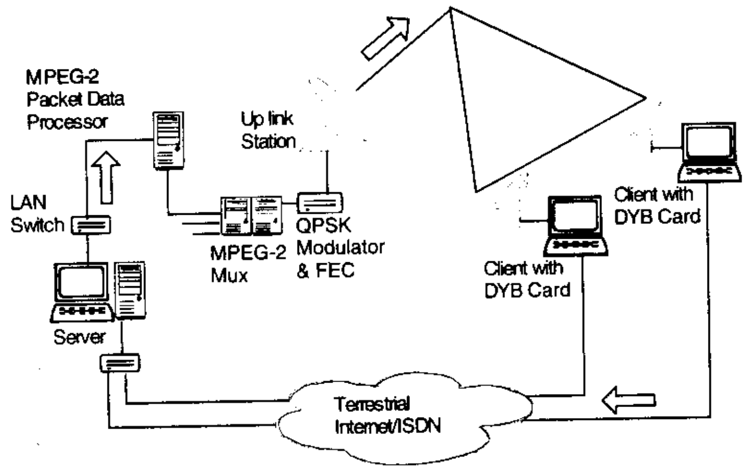

Figure 6.71 Typical configuration for providing DTH Internet delivery using DVB [6.204].

location. This presents challenges to the networking community, particularly as users become familiar with the benefits of high-speed connectivity. Along with an increased use of the Internet, there has been a revolution in TV transmission with the emergence of DVB. The same system may support a high-speed Internet and is being supported on a number of DVB satellite systems. A high-speed (6 to 34 Mb/s) simplex data transmission system may be built using a digital Low Noise Block (LNB) and standard TV antenna connected through an L-band coaxial cable to a satellite data receiver card in a PC (or LAN adaptor box). In many cases, a return link may be established using the available (standard dial-up modem) terrestrial infrastructure, providing the full-duplex communication required for the Internet service. Low-cost satellite return channels are also available. The overall system may provide low cost, high bandwidth Internet access to any location within link coverage of the DVB satellite service.

Data is already being sent across DVB networks using the MPEG-2 transport stream. A variety of proprietary encoding schemes are being used. Data transmission may be simplex or full duplex (using an interaction channel for the return) and may be unicast (point to point), multicast (one to many) or broadcast. Typical configuration for providing DTH Internet delivery using DVB is shown in Figure 6.71. In an effort to standardize services, the DVB specification suggests data may be sent using one of five profiles [6.204].

- *Data piping*—Where discrete pieces of data are delivered using containers to the destination.
- *Data streaming*—Where the data takes the form of a continuous stream that may be asynchronous (that is, without timing, as for Internet packet data), synchronous (that is, tied to a fixed-rate transmission clock, as for emulation of a synchronous

communication link) or synchronized (that is, tied through time stamps to the decoder clock).

- *MPE*—Based on DSM-CC and intended for providing LAN emulation to exchange packet data.
- *Data carousels*—Scheme for assembling datasets into a buffer that is played out in a cyclic manner (periodic transmission). The data sets may be of any format or type. The data is sent using fixed-sized DSM-CC sections.
- *Object carousels*—Resemble data carousels, but primarily intended for data broadcast services. The data sets are defined by the DVB network-independent protocol specification.

At the time DVB was being developed in Europe, a parallel program of standards and equipment development was also going on in the United States by the Advanced Television System Committee (ATSC). Among other things, ATSC adopted a different audio-coding standard and Vestigial Side Band (VSB) modulation. The United States has adopted a system based on ATSC DTV.

The MPEG-2 standards define how to format the various component parts of a multimedia program. They also define how these components are combined into a single synchronous transmission bit stream. The process of combining the streams is known as multiplexing. The multiplexed stream may be transmitted across a variety of links, such as the following

- Radio frequency links (UHF/VHF)
- Digital broadcast satellite links
- Cable TV networks
- Standard terrestrial communication links
- Microwave line of sight links (wireless)
- DSLs
- Packet/cell links (ATM, IP, IPv6, Ethernet)

Many of these formats are being standardized by the DVB project.

Each ES is input to an MPEG-2 processor that accumulates the data into a stream of PES packets. A PES packet may be a fixed or variable-sized block, with up to 65,536 bytes per block, and it includes a 6-byte protocol header. A PES is usually organized to contain an integral number of ES access units. The PES header starts with a 3-byte start code, followed by a 1-byte stream and 2-byte length field.

The MPEG-2 standard allows two forms of multiplexing.

MPEG Program Stream

This is a group of tightly coupled PES packets referenced to the same time base. Such streams are suitable for transmission in a relatively error-free environment and enable easy software pro-

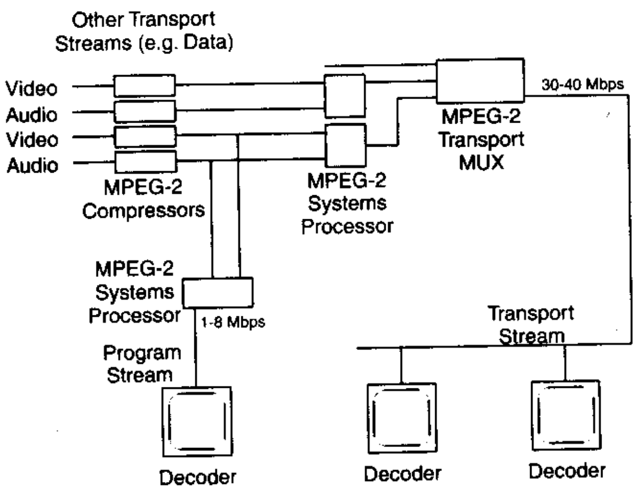

Figure 6.72 Combining ESs from encoders into a transport stream or a program stream [6.204].

cessing of the received data. This form of multiplexing is used for video playback and for some network applications.

MPEG Transport Stream

Each PES packet is broken into fixed-sized transport packets forming a general-purpose way of combining one or more streams, possibly with an independent time base. This is suitable for transmission in which there may be potential packet loss or corruption by noise, and/or where there is a need to send more than one program at a time. Combining ESs from encoders into a transport stream or a program stream is shown in Figure 6.72. The service information component on the transport stream is not shown. The program stream is widely used in digital video storage devices, and also where the video is reliably transmitted across a network. DVB uses the MPEG-2 transport stream over a wide variety of underlying networks. Because both the program stream and transport stream multiplex a set of PES inputs, interoperability between the two formats may be achieved at the PES level [6.205].

A transport stream consists of a sequence of fixed-size transport packets of 188 bytes. Each packet comprises 184 bytes of payload and a 4-byte header. One of the items in this 4-byte header is the 13-bit PID, which plays a key role in the operation of the transport stream. The format of the transport stream is described using Figure 6.73. This figure shows two ESs sent in the same MPEG-2 transport multiplex. Each packet is associated with a PES through the setting of the PID value in the packet header (the values of 64 and 51). The audio packets have been assigned PID64 and the video packets PID51. These are arbitrary, but different, values. As usual, there are more video than audio packets. Note that the two types of packets are not evenly spaced in time. The MPEG transport stream is not a time division multiplex because packets

Figure 6.73 Single program transport stream (audio and video PES) [6.204].

with any PID may be inserted into the transport stream at any time by the transport stream multiplexer. If no packets are available at the multiplexer, it inserts null packets to retain the specified transport stream bit rate.

Although the MPEG-transport stream may be directly used across a wide variety of media, it may also be used across a communication network. It is designed to be robust with short frames, each one being protected by a strong error correction mechanism. It is constructed to match the characteristics of the generic radio or cable channel and expects an uncorrected BER of better than 10^{-10}. The MPEG-2 transport stream is so called, to signify that it is the input to the transport layer in the ISO OSI seven-layer network reference model. MPEG-2 transport requires the underlying layer to identify the transport packets and to indicate in the transport packet header when a transport packet has been erroneously transmitted. The MPEG-transport stream packet size also corresponds to eight ATM cells.

Each MPEG-2 transport stream packet carries 184 bytes of payload data prefixed by a 4-byte (32-bit) header. The format of a transport stream packet is shown in Figure 6.74. The header has the following fields:

- The header starts with a well-known synchronization byte (8 bits).
- A set of three flag bits are used to indicate how the payload should be processed. The first flag indicates a transport error. The second flag indicates the start of payload. The third flag indicates a transport priority bit.
- The flags are followed by a 13-bit PID. This is used to identify the stream to which the packet belongs that was generated by the multiplexer. The PID allows the receiver to differentiate the stream to which each received packet belongs. Some PID values are predefined and are used to indicate various streams of control information. A packet with an unknown PID, or one with a PID that is not required by the receiver, is silently discarded.

The particular PID value is reserved to indicate that the packet is a null packet and is to be ignored by the receiver.

- The two scrambling control bits are used by conditional access procedures to encrypt the payload of some TS packets.
- Two adaptation field control bits may take four values:

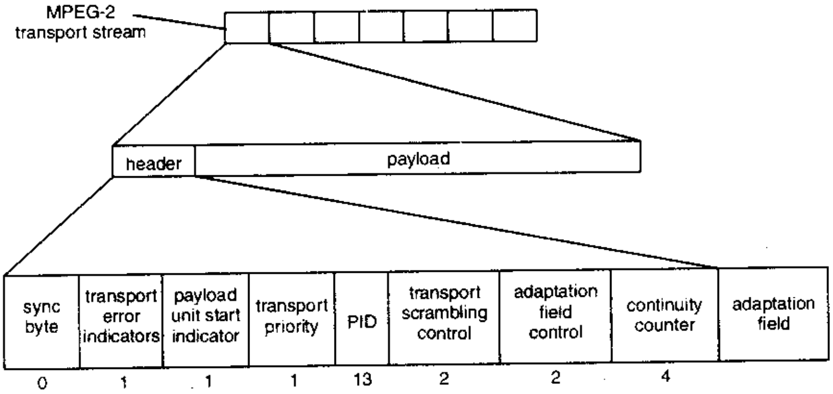

Figure 6.74 Format of a transport stream packet [6.204].

- 01 - No adaptation field, payload only
- 10 - Adaptation field only, no payload
- 11 - Adaptation field, followed by payload
- 00 - Reserved for future use.

- There is a half-byte continuity counter (4 bits).

Two options are possible for inserting PES data into the transport stream packet payload:

- The simple option, from both the encoder and receiver viewpoints, is to send only one PES (or a part of single PES) in a transport stream packet. This allows the transport stream packet header to indicate the start of the PES, but, because a PES packet may have an arbitrary length, it also requires the remainder of the transport stream packet to be padded, ensuring correct alignment of the PES with the start of transport stream packet.
- In general, a given PES packet spans several transport stream packets so that the majority of transport stream packets contain confirmation data in their payloads. When a PES packet is starting, however, the payload unit start indicator bit is set to "1," which means the first byte of the transport stream payload contains the first byte of the PES packet header. Only one PES packet can start in any single transport stream packet. The transport stream header also contains the PID so that the receiver can accept or reject PES packets at a high level without burdening the receiver with too much processing. This has an impact on start PES packets. MEPG PES mapping onto the MPEG-2 TS is shown in Figure 6.75.

DVB transmission may be received through a variety of equipment, as follows:

- A set-top-box
- An in-built decoder forming a part of a DTV set or a digital VCR

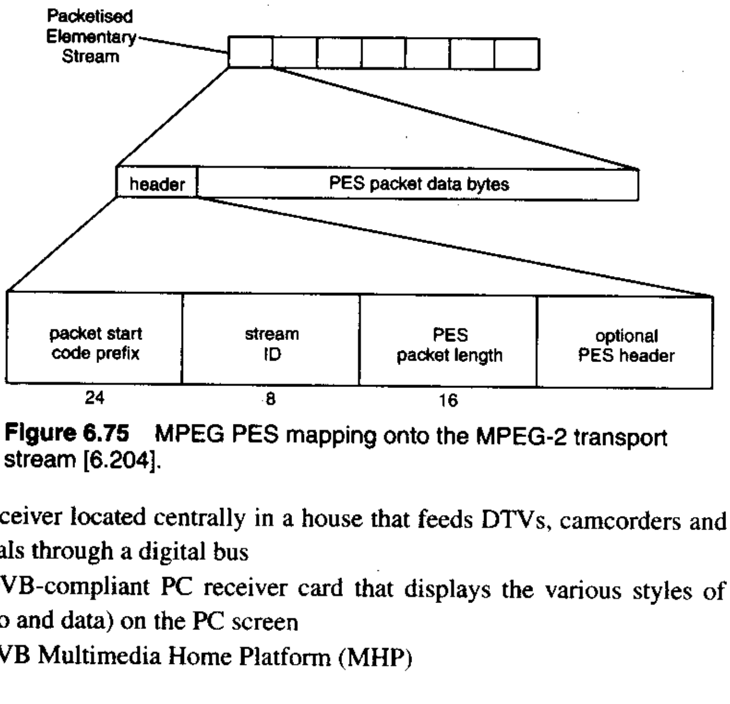

Figure 6.75 MPEG PES mapping onto the MPEG-2 transport stream [6.204].

- A receiver located centrally in a house that feeds DTVs, camcorders and so forth with signals through a digital bus
- A DVB-compliant PC receiver card that displays the various styles of content (TV, audio and data) on the PC screen
- A DVB Multimedia Home Platform (MHP)

The next generation of digital recorders will provide much greater functionality than existing VCRs. By transmitting information as MPEG metadata prior to each program, the digital recorder may itself determine whether the broadcast content should be recorded.

The MHP is defined to enable DVB receivers to be constructed in a common format and to provide common features and interfaces. DVB MHP uses the Java programming language to access an API, which gives access to DVB services. Three application profiles have been defined:

- Enhanced broadcast
- Interactive television
- Internet access

Broadband Multimedia Satellite Systems

DVB transmission by satellite (often known as DVB-S), defines a series of options for sending MPEG-transport stream packets across satellite links. The DVB-S standard requires the 188-byte (scrambled) transport packets to be protected by 16 bytes of RS coding. The resultant bit stream is then interleaved, and convolutional coding is applied. The level of coding may be selected by the service provider. The digital bit stream is then modulated using QPSK. A typical satellite channel has a 36-MHz bandwidth, which may support transmission at up to 35 to 40 Mb/s.

Multimedia is a term that may be applied to digital media-rich content, digital platforms and networks. The full range of multimedia services has been demonstrated over next-generation IP networks, leading to questioning as to the merits of introducing a unified ATM network. Types of multimedia services are the following:

- TV well suited to satellite
- VoD
- Electronic Program Guide (EPG)
- E-commerce services (shopping, gaming and so forth)
- Internet proxy services (selected Web and email)
- Internet access offered only by some current systems
- Games with surprising take up by some users

Most modern broadband systems also permit some level of interactivity. This is true of most current content (TV, radio and so forth), but is especially true of the new digital services, where users often want to sort, manipulate or participate based on the received content. Although it is possible to conceive a one-way Internet service (one-way routed Internet), most services will require some form of return channel to allow two way packet flow.

Multiplatform delivery is also a key concept for many people thinking of broadband multimedia. This is the ability to deliver content (digital video, Web pages and so forth) to a range of network devices, TV sets, PCs multimedia devices, Web-enabled telephones and wireless devices. There are five key players in a typical broadband multimedia system: content owners, middleware developers, service providers, network operators and customers.

Convergence is the term given to the perception that many platforms now have (or soon will have) common capabilities. One advantage of convergence is the ability to access the same information from a TV set as from a PC. In contrast to what is seen as a complexity of a PC, most potential broadband customers perceive the TV as the less difficult-to-use device. Nevertheless, to evolve the set-top-box will need to acquire much more sophistication and may never prove the ideal device for personal access.

The DVB Return Channel System by Satellite (DVB-RCS) was specified by an ad hoc ETSI technical group founded in 1999. This has tracked developments by key satellite operators and followed a number of pilot projects organized by the European Space Agency (ESA). The DVB-RCS system is specified in ETSI EN 301 790. This specifies a satellite terminal, sometimes known as a Satellite Interactive Terminal (SIT) or Return Channel Satellite Terminal (RCST), supporting a two-way DVB satellite system. The use of standard system components provides a simple approach and should reduce the time to market. The transmit capability uses a Multifrequency Time Division Multiple Access (MF-TDMA) scheme to share the capacity available for transmission by the user terminal. The data to be transported may be encapsulated in ATM cells, using ATM adaptation layer 5 (AAL-5) or may use a native IP encapsulation across MPEG-2 transport.

DVB-RCS terminals require a two-way microwave feed arrangement/antenna system to transmit and receive in the appropriate satellite frequency bands. These are typically connected by a cable (or group of cables) to an indoor unit. This unit could be a set-top-box within a network interface integrated in a PC peripheral or may be integrated in a PC expansion card. A key goal of DVB equipment suppliers is to reduce equipment costs.

Multimedia Home Platform

The MHP encompasses the peripherals and the interconnection of multimedia equipment through the in-home digital network. The MHP solution covers the whole set of technologies that are necessary to implement digital interactive multimedia in the home, including protocols, common API languages, interfaces and recommendations. At the beginning of 1996, the UNITEL-universal set-top box project was launched by the Program of the European Commission. The main aim of this project was to raise awareness of the benefits of developing a common platform for user-transparent access to the widest range of multimedia services. Promising progress has since been achieved toward the harmonization of what is now widely called the MHP. The MHP Launching Group was born from the UNITEL initiative in order to open the project to external parties through joint meetings. Key representatives of the High Level Strategy Group took part in this group, and this collaboration eventually led to the transfer of these activities to the DVB Project. Two DVB working groups were subsequently set up [6.206]:

- A commercially oriented group, DVB-MHP, to define the user and market requirements for enhanced and interactive broadcasting in the local cluster (including Internet access)
- A technical group, DVB-TAM (Technical Issues Associated with MHP), to work on the specification of the DVB API.

Different reference models have been defined for each MHP system currently in use. UNITEL used object-modeling tools to define the application classes and functionalities that would ultimately identify the hardware and software resources required by an MHP system. With this system, users would be able to access:

- Enhanced broadcasting services
- Interactive broadcasting services
- Internet services

This model offers system modularity through the use of key interfaces. These surfaces will be able to maintain the stability of MHP systems as they evolve, both in terms of hardware and software. Backward compatibility will be supported to the largest possible extent, for example, by using scalable applications.

The reference model consists of five layers [6.207].

- Application (content and script) and media (audio, video and subtitle) components
- Pipes and streams

- The API and native navigation/selection functions
- Platform/system software or middleware, including the interactive engine, the Run-time Engine (RTE) or virtual machine, the application manager and so forth
- Hardware and software resources and associated software.

Multimedia Car Platform

The Multimedia Car Platform (MCP) project is based on the results and achievements of the two predecessor projects, Multimedia Environments for Mobiles (MEMO) and Mobile Television and Innovative Receivers (MOTIVATE).

The major streams of the MCP projects are the following:

- Service definition and implementation based on user cases and service requirements for the car driver and passenger
- Specification for an open multimedia platform in the car integrating communication, entertainment and navigation
- Implementation of the first multimedia car terminal in the world
- Specification of the architecture of a hybrid network, including service handover and interoperability between different networks

The MEMO system architecture and protocol are used as the starting point for the work in MCP. However, although MEMO only provided specifications for DAB as a broadcast radio interface, in MCP this will be provided also for DVB-T [6.208]. An MCP network reference model is shown in Figure 6.76. MCP started work in January 2000. In a very short time, it attracted major interests of the car industry and has become one of the most important European manufacturers to get involved in MCP. MCP will encourage convergence among telecommunication, broadcasting and media, which actually have partly prohibitive cross regulations.

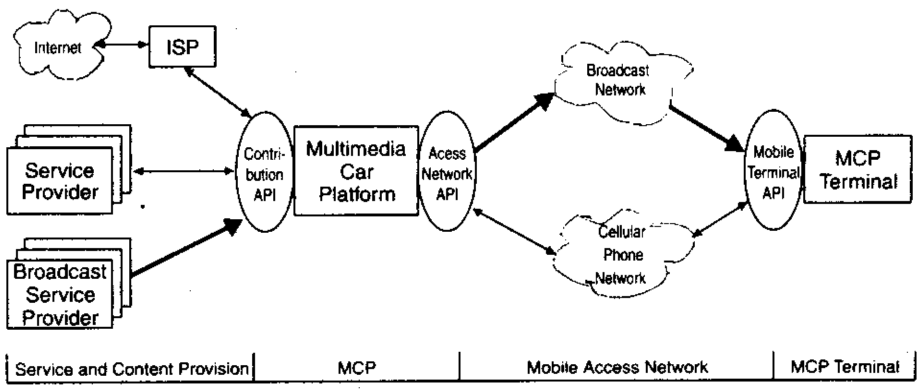

Figure 6.76 Multimedia car platform reference model [6.208].

MCP will actively promote changes in regulation in Europe to allow dynamic usage of time and frequency for data in broadcasting networks.

6.8 Digital Television Infrastructure for Interactive Multimedia Services

DTV technology appears commercially today in hybrid digital-analog systems, such as digital satellite and cable systems, and it serves as the default delivery mechanism for HDTV. All digital sets, such as HDTV, can display higher resolution of digital format and do not require additional external conversion equipment. The digital format for TV broadcast transport has several key advantages over analog transport. For service operators, the key benefit is the high-transport-efficient digital compression that packs five or more times as many channels in a given distribution network bandwidth. This, in turn, increases the operator's revenue potential by delivering more content and pay-per-view events, including Near Video-on-Demand (NVoD) movies with multiple, closely spaced start times. End-users have a larger program selection with spaced start times. End-users have a larger program selection with CD-quality audio and better picture quality potential even when viewed in a hybrid setup with analog TV sets [6.209].

Figure 6.77 shows an example of a satellite-based DTV broadcast system. However, most of the discussion that follows is independent of the physical distribution network and applies to cable and over-the-air digital transmission as well.

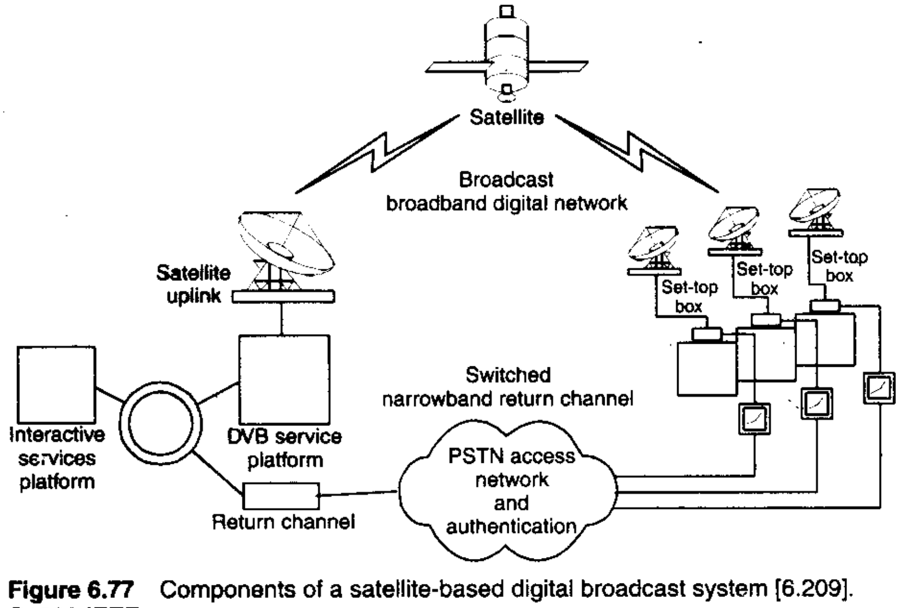

Figure 6.77 Components of a satellite-based digital broadcast system [6.209]. ©1998 IEEE.

A digital video broadcast network distributes audio and video streams to subscribers using a transport protocol. In standard-based implementations, the MPEG-2 transport stream carries digital data across the broadcast network. The MPEG-2 transport structure may contain multiple video and audio channels, as well as private data. MPEG PIDs uniquely identify all program components. A typical digital set-top box contains a control microprocessor and memory, a network interface and tuner, demodulator for cable and satellite, a transport stream demultiplexer and MPEG audio and video decoders. A set-top box also implements user interface components such as the remote-control input and on-screen graphical display capability for output, which are used for controls, and the electronic programming guide. Where appropriate, the conditional access keys distributed to authorized subscribers are used to decrypt the encrypted content in the set-top boxes. Similar functionality can also be packetized into PC-compatible interface cards to let PCs receive digital video, audio and data from the broadcast network. Within the constraints of their service subscriptions, individual users can watch any channel by tuning into the appropriate program within the broadcast multiplex. The individual set-top box that resides on the user's premises handles the channel tuning.

From the data delivery point of view, the DTV infrastructure provides a broadband digital distribution network, data transport protocols and digital terminals (set-top decoders) on the user's premises. As such it provides a powerful platform for delivering information and data services that not only enrich, but fundamentally transform, the television-viewing experience. DTV systems always provide a one-way broadcast path for distributing digital video. Optionally, a return communication link can be provided to allow the upstream flow of data from users to the service center. The return channel often supports the impulse buying of pay-per-view and NVoD events. The return channel is usually implemented through a narrowband communication link such as a PSTN, or an ISDN. Cable systems with two-way-enabled plants can implement a return channel across the cable infrastructure. Because both cable and satellite DTV systems use the same data transport mechanism and protocols, the MPEG-2 transport stream, the physical nature of the underlying distribution network is transparent to data services.

Data broadband technology enables DTV service providers to enhance their customers' television-viewing experience by providing a wide range of interactive services as an incremental add-on to the DTV broadcast infrastructure. Depending on the underlying DVB system infrastructure, the following classes of interactive services are possible:

- Broadcast-only interactive services
- Broadcast with a batch return channel
- Broadcast with an online return channel

Table 6.1 summarizes types of broadcast-based services that a DTV infrastructure 'ith different return channel options can deploy. Note that user-level capability and interactivity is a function of the network and connectivity infrastructure. An important challenge facing DTV service designers lies in dividing data services that operate in the most common, broadcast-only

Table 6.1 Interactive data services as a function of underlying infrastructure capability [6.209].

Functionality	Digital broadcast (one to many) downstream				Point-to-point
	One-way plant (satellite or cable)			Two-way plant (cable only)	
	No return	Polled return (PSTN)	Real-time return (PSTN)	Cable return (real-time)	
Network	Broadcast One-way	Broadcast Polled return	Broadcast Phone return, dial-up	Broadcast Two-way HFC	Switched Two-way FTTC, ATM
Interactivity	Local	One-way (user response)	Two-way	Two-way	Two-way
User-level function	Browse (view interactively)	Browse plus batch transaction	Browse plus real-time transaction	Browse plus real-time transaction	Full service

FTTC: Fiber To The Curb

environment and that scale up in user-level functions with the increased capability of the underlying DTV infrastructure, when available.

Interactive Broadcast Data (IDB) Services

Interactive broadcast data services can be broadly categorized as follows:

- *IDB services*—Provide primarily data-only channels, with optional background audio. When a user selects such a service, a data-only screen displays. The user may use hot-spot or hyperline-style mechanisms.
- *Interactive Video Broadcast (IVB) services*—Combine video and data channels to provide an enhanced TV-viewing experience. Service delivery and user experience of combined video-data services can be further categorized in terms of their temporal relationship:
  - *Unsynchronized video and data*—Video and data may be typically related or unrelated. Common examples include a simple interactive icon overlay, a partial data overlay such as a ticket display or a data screen with a video (broadcast) insert.
  - *Synchronized video and data*—In this mode, video and data are both typically related and authorized to be synchronized at playback.

Table 6.2 shows a sampling of interactive services.

Table 6.2 Classification of sample services [6.209].

Service	Broadcast (no return channel)	Polled return channel	Real-time return
Primary user-level service capability	Browse (local interactivity)	N/A	N/A
Broadcast Video Services			
Broadcast video	Tune	Tune	Tune
Electronic program guide	View, tune to selection	View, tune to selection	View, tune to selection
Impulse PPV, NVoD	View (order by phone)	View, order (smart card log)	View, order
IDB Services			
Information services	Browse (data carousel)	Browse and acknowledge	Browse and request
Games	Download and play	Download and play (delayed comparison of scores)	Download and play, real-time comparison of scores, multiplayer
Home shopping	Browse (order by phone)	Browse, order (delayed confirmation)	Browse, order
IVB Services			
Enhanced program information	Additional information broadcast, synchronized on current video program	Additional information broadcast, with delayed requests	Fully interactive, ability to request additional information
Interactive advertising	Browse (service information)	Browse, order coupons, brochures, goods (delayed confirmation)	Browse, order online
Play-along programming	Play along, keeps local score	Play along, keeps local score, delayed comparison of scores	Play along, local and networked scoring in real time

Table 6.2 Classification of sample services [6.209]. (Continued)

Service	Broadcast (no return channel)	Polled return channel	Real-time return
Online Services			
Email, forums	Receipt only	Receipt with delayed reply, post	Full interactive receive and replay
Internet access	Intranet broadcast	N/A	Full service
VoD	Not supported	Not supported	Fully interactive

Many DTV systems use an intermittent batch return channel for billing of Impulse Pay-Per-View (IPPV) events. In this approach, the return channel is usually implemented using a telephone line with a low-speed, dial-up modem. At specified intervals, the control billing system polls individual set-top boxes to retrieve the accumulated information. Optionally, the set-top box might dial up the central system when its local event storage almost fills up or the IPPV viewing credit nearly runs out.

Data Carousel Concept

To support low-end set-top boxes with small amounts of application memory and no additional storage capacity, the datastreams are broadcast cyclically on the DTV network. In effect, the DTV network acts as a large serial disk for storage. This approach gives rise to what is known as the data carousel. The approach allows clients with a local caching capability to find the necessary data and code on the network at any time, with the worst-case access latency equal to the carousel cycle duration. Figure 6.78 shows the basic layout of a data carousel. The carousel datastream consists of an indexing and naming mechanism to locate objects within the data carousel, application code to download to the receiver when the user tunes into the carousel data channel and application data objects that the user terminal retrieves at run time in response of the user's interactive requests.

A DTV network provides an ideal platform for distributing data broadcasts in a carousel fashion. The interactive data services (carousel) are a multiplexed MPEG-2 transport stream. From the management and distribution points of view, data service files can be handled in exactly the same manner as any other DTV stored content (such as NVoD). This lowers system acquisition and operation costs for interactive services because the familiar service center equipment and procedures for TV signal delivery, such as the scheduler and NVoD server, also distribute data services.

For more detailed information on these topics, the references [6.210, 6.211, 6.212] are recommended as research papers dealing with some theoretical concepts of data broadcast. A

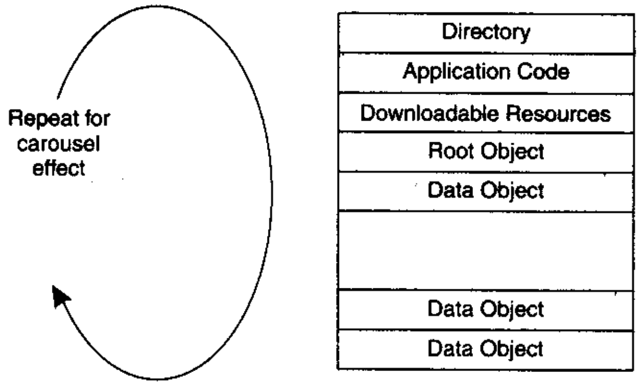

Figure 6.78 Structural layout of a data carousel [6.209]. ©1998 IEEE.

detailed treatment of DTV components and infrastructure can be found in [6.213]. The reference [6.214] deals with user-interaction issues.

6.9 Concluding Remarks

Research on telecommunication networks is focused on post-ISDN architectures and capabilities, such as an integrated packet network and BISDN. The economics and flexibility of integrated networks make them very attractive, and packet network architectures have the potential for realizing these advantages. However, the effective integration of speech and other signals, such as graphics, image and video, into an IPN can rearrange network design priorities. Although processing speeds will continue to increase, it will also be necessary to minimize the nodal per-packet processing requirements imposed by the network design. This is a motivation for new switching concepts like fast packet switching and ATM. Data signals must generally be received error free in order to be useful, but the inherent structure of speech and image signals and the way in which they are perceived allows for some loss of information without significant quality impairment. This presents the possibility of purposely discarding limited information to achieve some other goal, such as the control of temporary congestion.

Layered coding refers to coding a video into a base layer and one or several enhancement layers. The base layer provides a low but acceptable level of quality, and each additional enhancement layer will incrementally improve the quality. By itself, layered coding is a way to enable users with different bandwidth capacities or decoding powers to access the same video at different quality levels.

When transmitting a video signal, residual errors are inevitable regardless of the error-resilience and channel-coding methods used. Thus, ways of mitigating these errors have to be devised. A number of approaches have been proposed in the literature toward the error concealment. Such approaches can be classified into spatial and temporal domain approaches. Compared to coders that are optimized for coding efficiency, error-resilient coders typically are less efficient in that they use more bits to obtain the same video quality in the absence of any transmission errors. These extra bits are called redundancy bits, and they are introduced to enhance the video quality when the bit stream is corrupted by transmission errors. The design goal in error-resilient coders is to achieve a maximum gain in error resilience with the smallest amount of redundancy.

The main challenge in designing a multimedia application across communication networks is how to deliver multimedia streams to users with minimal replay jitters. To diminish the impact of the video quality due to the delay jitter and available network resources (for example, bandwidth and buffers), traffic shaping and SRC are qualified candidates at two different system levels. Traffic shaping is a transport layer approach, and SRC is a compression layer approach. The SRC approach is compressed according to the application's requirement and available network resources.

Streaming video across the Internet faces many technological as well as business challenges, and new codecs, protocols, players and subsystems have been developed to address them. Since its introduction in the early 1990s, the concept of streaming media has experienced a dramatic growth and transformation from a novel technology into one of the mainstream manners in which people experience the Internet today. The concept of streaming media comes at a time when basic multimedia technologies have already established themselves on desktop PCs. Streaming media is a technology that enabled the user to experience a multimedia presentation on-the-fly while it was being downloaded from the Internet.

The provision that bandwidth on demand with strict QoS guarantees is a fundamental property of ATM networks that makes them especially suitable for carrying real-time multimedia traffic. Statistical multiplexing of VBR connections within the backbone network allows effective aggregation and capacity engineering. Several statistical aggregation mechanisms (such as those based on equivalent bandwidth computation) can be used to dimension appropriately the capacity for backbone transport. Scalability is thus possible because per-connection explicit bandwidth renegotiation is not needed within the backbone. Bandwidth renegotiation frequency is high only with the access nodes, where the expected number of high bit-rate multimedia connections is not too large. In addition, the flexibility of soft-QoS control allows applications to specify a wide range of QoS expectations, from best effort to strict guarantees. Because only transport-level parameters are renegotiated, the framework can be applied independently of the network layer. The soft-QoS specification can be used independently by various network technologies to ensure that users' end-to-end service expectations are met.

Anticipating that multimedia across IP will be one of the major driving forces behind the emerging broadband communications, addressing the challenges facing the delivery of multimedia applications across IP is of great importance. In order for the Internet to allow applications to request network packet delivery characteristics according to their needs, sources are expected to declare the offered traffic characteristics. Admission control rules have to be applied to ensure that requests are accepted only if sufficient network resources are available. Moreover, service-specific policing actions have to be employed within the network to ensure that nonconforming data flows do not affect the QoS commitments for already-active data flows. One generic framework that addresses both the video-coding and networking challenges associated with Internet video is scalability. Any scalable Internet video-coding solution has to enable a very simple and flexible streaming framework. The fine-grained scalable framework strikes a good balance between coding efficiency and scalability while maintaining a very flexible and simple video-coding structure.

With the advent of common uses of the Internet, the demands for real-time and low-rate VoIP applications are growing rapidly. Because the delivery of packets is not guaranteed in the IP networks, it is necessary to deal with the audible artifacts, which are caused by burst packet losses. Packet loss seriously degrades the speech quality of the analysis-by-synthesis coders because the loss parameters not only affect the current speech frame, but also produce the so-called error-propagation problem resulting from corrupted filter memory. This packet-loss problem can be solved by using different model parameters.

DSL technology offers unprecedented scalability for interactive video services. It is the basis for the point-to-point architecture that is the key to providing a combination of interactive video and broadcast services. The implementation of video services is a high priority for telecom providers. Delivering voice services across DSLs offers a lucrative opportunity for both established and emerging services.

Data broadcasting in support of multimedia applications requires efficient use of bandwidth resources in order to maximize the availability of playout content. From the data delivery point of view, the DTV infrastructure provides a broadband digital distribution network, data transport protocols and digital terminals on the user premises.