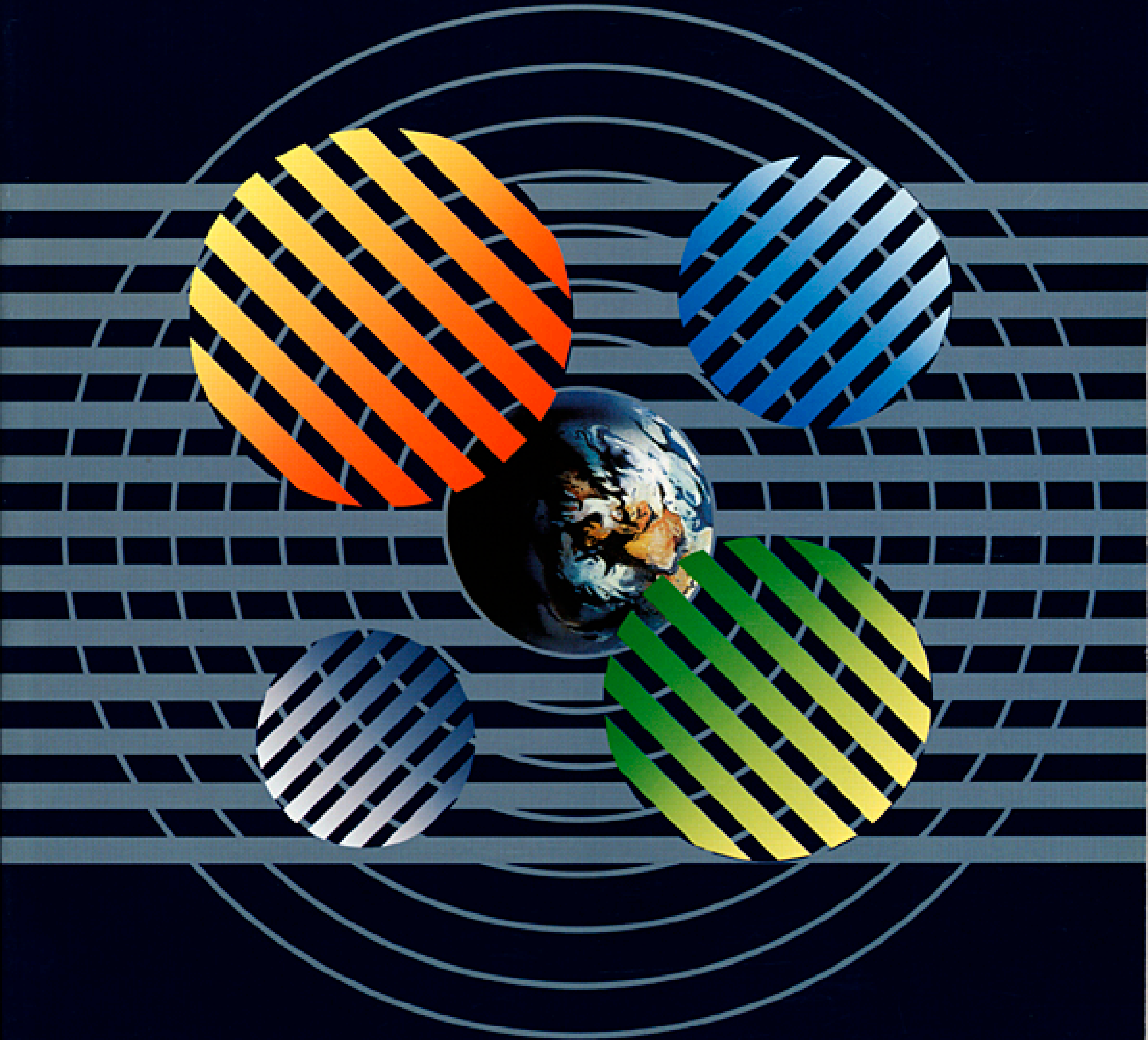
Telektronikk 2.94

Broadcasting



FEATURE:

Editorial, *Tore Øvensen* **3**

The technical development of broadcasting in recent years and towards year 2000 – Special emphasis on the conditions in Norway, *Tore Øvensen* **5**

Perceptual audio coding for digital sound broadcasting, *Erik Bernatek* **13**

Digital Sound Broadcasting – Transmission Systems and Terrestrial Networks, *Eivind Vågslid Skjæveland* **20**

Digital Audio Broadcasting: Terrestrial transmission experiments in Norway, *Rolf Møkleby* **26**

Radio production – evolution towards an all digital environment, *Are Nundal* **31**

Television production – evolution towards all digital systems, *Per Bøhler* **35**

Digital Television based on MPEG2 – Activity in Norway, *Gisle Bjøntegaard* **44**

HD-DIVINE – a Scandinavian Digital Television System, *Per Appelquist* **49**

Optical Routing Systems in Radio and Television Production Centres, *Bjørn M Elgstøen* **55**

Olympic Host Broadcaster Operations, *Arild Hellgren and Helene Amlie* **65**

STATUS:

International research and standardisation activities in telecommunication: Introduction, *Endre Skolt* 7 3

The TINA Consortium, *Tom Handegård* **74**

Telecommunications languages and methods, *Arve Meisingset* **81**

Message Handling Systems, *Geir Thorud* **86**

Security, *Sverre Walseth* **89**

Telektronikk

Volume 90 No. 2 - 1994
ISSN 0085-7130

Editor:
Ola Espvik
Tel. + 47 63 80 98 83

Status section editor:
Endre Skolt
Tel. + 47 63 80 92 11

Editorial assistant:
Gunhild Luke
Tel. + 47 63 80 91 52

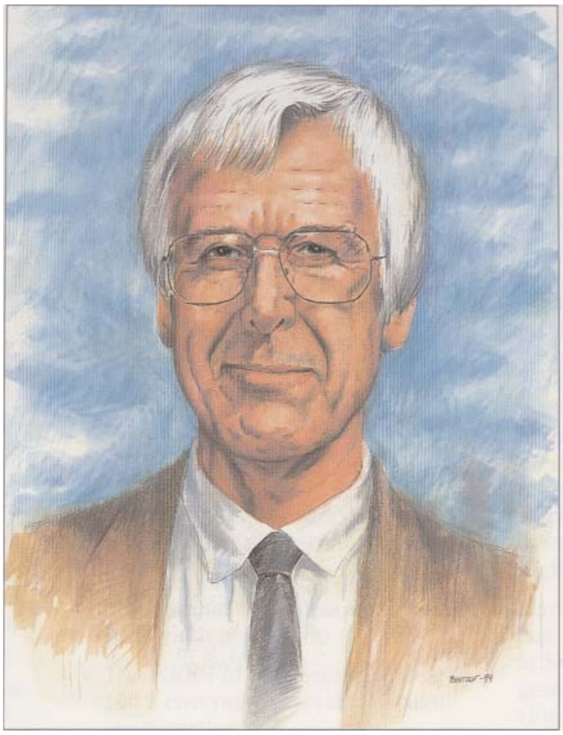
Editorial office:
Telektronikk
Norwegian Telecom Research
P.O. Box 83
N-2007 Kjeller, Norway

Graphic design:
Design Consult AS

Layout and illustrations:
Gunhild Luke, Britt Kjus, Åse Aardal
Norwegian Telecom Research

Guest editorial

BY TORE ØVENSEN



Sound broadcasting in Norway started as soon as technology made it possible.

For most of this time, from 60 years back and up to the 1980's, NRK (Norwegian Broadcasting Corporation) has been the only public broadcaster in Norway both for Radio and Television. The Norwegian Telecom has been responsible for the transmission network.

It has been an enormous task and very expensive to establish the NRK national Radio and TV network. This is due to the difficult topography with narrow valleys and fiords and a scattered population.

Today, NRK provides one national TV channel, three national FM radio channels and an international medium wave and short wave service for Norwegians abroad.

The TV network alone consists of 45 large VHF main transmitters and more than 2,500 repeaters for secondary coverage, giving a total coverage of about 99 % of the population.

The NRK main production centre for TV, Radio P2 and Radio Norway International is in Oslo, while the main production centre for Radio P1 and P3 is in Trondheim. A total of 17 local production centres all over the country contributes to both local programmes and the national programmes for Radio and Television.

All programmes from NRK are also transmitted in the D-MAC system in the Nordic broadcasting satellite TELE-X. This duplication of transmission is mainly aimed at covering the Spitzbergen islands and the Norwegian oil installations in the North Sea.

Today, NRK is not alone on the air. TV2, a commercial national Public Broadcaster, started television transmission in 1992. Most of the TV2 terrestrial networks are based on programme feeding from a satellite (Intelsat) in the D2-MAC system. TV2 is therefore also available at Spitzbergen and in the North Sea, and the satellite transmission gives also supplementary coverage of the Norwegian mainland.

P4, a commercial FM Radio channel with national coverage, started transmission in 1993. In addition, a large quantity of small local FM Radio stations have been established all over the country during the last 10 years.

The national commercial programmes will never cover more than about 95 % of the population – the price for reaching full coverage is too high.

The broadcasting technique up to now has mainly been based on analogue technique, both in the studios and on the transmission side.

During the next few years a radical change to all-digital systems will take place. Programme production, both in Radio and TV, is already partly digital and transmission systems based on bit compression techniques will gradually replace FM sound transmissions and PAL TV transmissions as we approach the turn of the century.

The introduction of new digital technology will also in the longer term give a change in picture format from the aspect ratio 4:3 to 16:9. This broad picture format was primarily developed for the next generation TV, HDTV – High Definition Television. HDTV production facilities are already available, but the introduction of the system for the public will most probably be put off until the beginning of the next century. The bottleneck is the delayed development of flat large screen HDTV receivers for the consumer market.

The developments mentioned above are a great technical challenge and the following pages will give you more detailed information of the new fascinating techniques.

You will sooner or later be familiar with the newcomers:

- DAB – Digital Audio Broadcasting
- DVB – Digital Video Broadcasting
- HDTV – High Definition Television.

This edition of Telektronikk will also deal with an important event for the introduction of new TV production and transmission technology, the XVII Olympic Winter Games at Lillehammer in February of this year.

NRK, being the host broadcaster, decided to introduce as far as possible the digital 625 line production standard, and this event was a breakthrough for a large scale use of this standard and for the transmission of the digital signals on fibre optic cables in the Lillehammer region.

Tore Øvensen

The technical development of broadcasting in recent years and towards year 2000

– Special emphasis on the conditions in Norway

BY TORE ØVENSEN

1 Introduction

The general conditions for broadcasting in Norway are rather extreme:

- The ground conductivity is very low due to the mountains that consist mainly of the hard rock types gneiss and granite. The range of the ground waves for long and medium wave transmitters are therefore low, and very high power transmitters are needed to give satisfactory coverage.
- VHF and UHF radio and television transmission, which is nearly a line of sight transmission, must fight against high mountains and steep valleys. In some parts of the country, fiords at the bottom of the valleys give reflections from the water surface causing special problems.
- The high latitude causes low elevation angles for satellite broadcasting. The elevation angle is about 24 degrees in the south and about 10 degrees in the north of the main-land resulting in shadows in east-west going valleys. Large buildings and trees can also cause trouble.
- The radio link network for feeding transmitters and for programme contribution has an extension of about 2,000 km in the south-north direction. This very long transmission path gives special requirements for reliability and signal quality.

All these extreme conditions have been a great challenge when establishing the infrastructure for broadcasting in Norway. Large investments have been necessary to give satisfactory coverage for the population both for the Radio and Television networks.

2 Sound Broadcasting

2.1 Long and medium wave network

Until 1975 the NRK sound broadcasting network comprised one radio programme (P1) using AM long and medium wave transmitters and partly FM transmitters. In 1978 most of the existing AM network was discontinued according to the new international plan for the LF/MF broadcasting bands.

This plan assigned four very high power channels (1200 kW) for Norway. But this plan has been only partly realized in our country. One 1200 kW medium-wave station has been established by NRK at Kvitsøy near Stavanger, mainly for serving the Norwegian merchant fleet, the fisheries in the North Sea and the Norwegian Sea, and Norwegians living and travelling in Europe. Two more of the assigned frequencies are in use, transmitting at reduced power, Vigra Ålesund at 100 kW/MF and Kløfta/Oslo at 200 kW/LF. These stations are old and will probably never be replaced by 1200 kW transmitters for the following reasons:

- The investments needed are very high and so are the running costs.
- The technical quality is limited to mono sound and a maximum audio bandwidth of 4.5 kHz.
- The NRK FM networks give nearly 100 % coverage, at least for stationary reception.

The existing AM network is far from sufficient for full coverage of the country, but has some value as supplementary coverage for mobile reception, especially in West Norway, where the topography is a great problem for FM reception.

In the future, after the introduction of terrestrial digital audio broadcasting (DAB), some of our AM transmitters will still be important for the fishing fleet and for the coverage of Europe. If DAB transmission by satellite will be introduced at the beginning of the next century, AM broadcasting at long and medium waves could be superfluous.

2.2 International broadcasting on short wave

The short wave bands allocated to broadcasting have been and will continue to be very important for international sound broadcasting. The international frequency conference, WARC 92, allocated more frequency bands for this service, but not sufficient to cover the requirements for channels. Changing to single sideband transmission (SSB) will reduce the congestion in the bands and was strongly recommended by WARC 92. But SSB means investments in new transmitters and will take a long time, especially in developing countries, because of the high investment costs.

During the last 10 to 15 years NRK has built two new short wave stations at Kvitsøy/Stavanger, and Sveio/Haugesund on the Norwegian west coast (Figure 1). These two stations are complementary and jointly they cover all relevant transmission directions. The stations are equipped with 500 kW transmitters and are also able to transmit in the SSB modus. NRK has already used so-called compatible SSB transmission for some years in the highest frequency bands. In addition to the upper sideband, the carrier is transmitted with a 6 dB power reduction. After a very long transition period, lasting into the next century, the final modulation system will be introduced. The carrier will then be reduced by 12 dB, giving the full benefit of the system, which is interference reduction and power savings at the transmitter stations.

Figure 1 NRK short wave station at Kvitsøy/Stavanger

This world wide short wave programme service will continue because short wave transmissions cannot be fully replaced by other systems. The main advantage is that the programmes can be received as far away as the opposite side of the globe with very small and rather cheap receivers.

2.3 The FM VHF network

FM broadcasting of sound programmes in the band 87.5 MHz – 100 MHz, band II, started in Norway (NRK) towards the end of the 1950–1960 decade. This new service revolutionized the technical quality of sound broadcasting and qualified to the term Hi-Fi (high fidelity), when using a stationary antenna and a top quality home receiver. This transmission was matching in quality the new vinyl-based LP records. The ability of rendering the full quality of an LP record came, however, when FM stereo was on the air – in Norway about 1970.

The introduction of the pilot-tone stereo system gave a considerable reduction of the service area of the FM transmitters. At least 6 dB higher field strength was necessary, and more repeaters (slaves) were needed in shadow areas.

In addition to about 50 high power main transmitter stations, more than 1,600 low power repeaters were built giving nearly 100 % coverage for NRK P1 and P2 in the year 1990, for stationary reception using top roof antennas.

The pilot-tone stereo system also had a serious impact on the quality of mobile reception giving more distortion when receiving multipath signals.

The frequency conference RARC-FM 1984 extended band II in Europe from 100 MHz to 108 MHz, giving room for many new services. In Norway the new FM plan comprises four national networks (P1–P4) and hundreds of local low power transmitters, mainly intended to cover all municipalities in the country with local programmes.

The third NRK programme, P3, and the commercial P4 were launched in the autumn of 1993. P3 and P4 will not be given full coverage for the population. For economic reasons the coverage will be limited to about 90 % – 94 %. Full coverage for these late coming FM networks would be a waste of money, since the introduction of a DAB network probably will start within a few years.

2.3.1 RDS – Radio Data System

The density of FM transmissions is very high in most European countries, and tuning to a given programme is very difficult, particularly for listeners using portable or car radios. During the last 10 years new technologies have offered the possibility of adding to FM sound programmes auxiliary data signals which give methods for identifying the transmission and thereby facilitate automated tuning functions. This is RDS. A low bitrate signal (1187.5 bit/s) is amplitude modulated on a 57 kHz subcarrier, locked to the 19 kHz pilot-tone, without causing quality impairment of the programme signal (Figure 2).

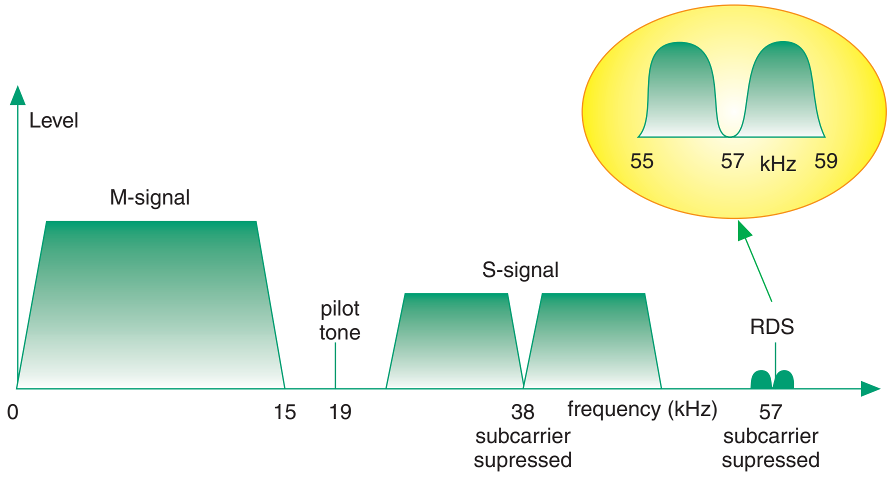


Figure 2 Base band spectrum including RDS, FM sound transmission

The RDS main features are

- *programme identification,* which labels the transmission station
- *programme service name,* which gives the name of the programme provider in a receiver display
- *alternative frequencies.* A list of alternative frequencies are transmitted giving the receivers the possibility to check the signal strength and phase distortion of the listed frequencies and choose the best one.

These features are called static information because they are always present during transmission.

Since the Norwegian sound programme networks consist of hundreds of low power repeaters for coverage of shadows in the service areas, car listening on FM stations is very inconvenient. In some parts of the country retuning of the car receiver is necessary within short distances, and as already mentioned multipath distortion is high in the same areas. NRK considered therefore the implementation of RDS as an urgent matter, and the system came into operation for P1 and P2 in 1989, as soon as RDS car receivers were on the market.

From 1993 NRK has implemented, on a small scale, the transmission of dynamic RDS, which means labelling the different programme types, for instance sports, drama, traffic information. This service is at present limited to local traffic information, when the P1 network is split up for transmission of local programmes from the 17 regional programme centres. Further development of the programme type coding has not yet been decided.

2.4 Digital Audio Broadcasting, DAB

There is an increasing demand for better broadcast sound quality because the quality of recorded sound has improved considerably through the development of digital recording techniques (CD – compact disc and DAT – digital audio tape). The existing FM sound broadcasting system cannot fulfil such a demand. This is particularly true under mobile reception conditions. The increasing demand for new programme services has also required more intensive utilization of the frequency spectrum for FM sound broadcasting (band II). In addition, the extensive use of compression of the audio signals has increased the average modulation, and thus the sideband energy levels.

The result is more interference problems and consequently difficulties for the listeners to obtain satisfactory reception.

To take full advantage of the audio quality that has become possible through the digital production and recording technique, EBU members have been working for many years on an advanced digital sound broadcasting system called DAB, intended for fixed, portable and mobile reception. Tests and demonstrations since 1988, using advanced modulation and source coding techniques, have shown that this new system provides very high quality even under severe multipath conditions. It offers this great advantage while also making extremely efficient use of the frequency spectrum.

The main reason for this high spectrum efficiency is the very high bit rate reduction in the source coding, and that the system makes it possible to use single frequency networks. As an example, the 20.5 MHz width of band II (87.5 – 108 MHz) can accommodate about five national stereo programmes with FM technique and at least 10 times more programme channels using DAB single frequency networks. This estimation is based on the final system concept, which implies a time division and frequency division multiplex of 5 to 6 stereo programmes within a 1.5 MHz frequency block.

The DAB system can be used for terrestrial and satellite broadcasting. The final development of the system was taken over by the Eureka 147 DAB project, working in close co-operation with the EBU. The chosen source coding, MUSICAM (ISO/MPEG Layer II), and the COFDM modulation system will be described later in this issue.

2.4.1 Frequency bands for DAB

The ITU WARC 92 adopted the band 1452–1492 MHz primarily for BSS sound (Broadcasting Satellite Service Sound), but the band may also be used for terrestrial DAB. This band will be fully available for BSS as late as the year 2007. For this reason European broadcasters must plan for a terrestrial implementation of DAB about 10 years earlier. The 1452–1492 MHz band can be used terrestrially in some countries much earlier, but this is not favourable because the spacing between transmitter sites must not exceed about 10 km in a single frequency network, due to the doppler-frequency shift at car reception. Some countries are, however, making plans for using this frequency band for local services, for instance covering a large city.

It has been found from theoretical and practical studies that frequencies in band II (87.5–108 MHz) and band III (174–230 MHz) are particularly suitable for terrestrial DAB transmission. Within these frequency bands the existing broadcasting stations (FM and TV) with a distance of 60–80 km is usable, and large investments in new transmitter site infrastructure is not necessary. Band I (47–68 MHz) is also suitable, but the level of man-made noise is higher, and significant sporadic-E propagation can give periods with increased interference.

2.4.2 Implementation of DAB in Norway

Eureka 147 has given EBU members the offer of purchasing DAB prototype equipment for test transmissions. This offer was the incentive of the NRK management to apply for a licence for the operation of a test network in collaboration with Norwegian Telecom and P4.

The licence was given and a provisional frequency band allocated at about 235 MHz for operating a single frequency network, consisting of three transmitter sites in the Oslo area. The network is equipped for transmission of one frequency block of 1.5 MHz with a capacity of four stereo programmes. This experimental transmission started in April 1994.

NRK is prepared to give high priority to the implementation of DAB in its sound programme network. The point in time when a regular DAB programme service can start depends on mass production of receivers and consequently of the development of DAB transmissions in other European countries. A realistic date would be 1997. The total cost of a single frequency network covering Norway will determine the implementation pace.

The existing sound programmes have to be transmitted both on FM and DAB probably for 10 years or more. This simulcast is necessary to give a reasonable time for the listeners to replace old receivers. The long transition period is a concern for the broadcasters because of considerable extra running expenses.

Implementation of additional national programmes is not yet decided.

3 Television

The existing colour television systems, NTSC, PAL, and SECAM have been the only production and transmission systems for more than 25 years. These systems were developed for compatibility reasons from the preceding black and white 525/60 and 625/50 systems that are nearly 50 years old. The colour systems are all so-called composite systems where the colour information is carried by a subcarrier within the luminance channel bandwidth. These systems have considerable quality limitations compared with what is possible with the technology available today. But one must remember that the main philosophy behind the development of these systems was cheap receivers, and so far back in time complexity meant high cost. With the enormous progress in semiconductor circuitry in recent years there is no special cost advantage any longer in simplifying receiver processing.

A fundamental limitation is the scanning process. A picture frame consists of two line-interlaced fields. Moving images can therefore give rise to interline twittering and line crawling. This is most visible in the 625 line system where the frame rate is only 25 Hz. Interlacing was chosen for reducing frame flickering and the necessary transmission bandwidth. The composite signal format in colour television causes intermodulation effects, cross colour and cross luminance, which is more or less visible dependent on the type of picture.

During the first 10 to 15 years of the colour TV transmission these shortcomings of the systems have not been the limiting factors of the picture quality delivered to the homes over the terrestrial networks:

- On the production side there has been enormous progress in technical quality of the production equipment during the last 10 years. The colour rendering and the noise level of the electronic cameras have been improved year by year. The last family of cameras, using CCD (charge couple device) instead of pick up tubes, is able to deliver a picture quality much higher than can be rendered after PAL coding.
- Most of the programme production is based on videotape recording and editing. The recording technique has also been improved, especially in later years, by replacing composite recorders with analogue component

recorders. A first generation and very expensive digital component recorder has also been on the market for some years. But viable, reasonable priced digital component recording machines were just recently released for sale. The video recordings at Lillehammer Olympic Games were based on such recorders. The digital signal format is in accordance with recommendation ITU-R BT.601 and 656 (former CCIR). In the component recording technique the luminance and colour signals are recorded on separate tracks with no cross effects. When also the succeeding editing process is based on component technique, the quality of the source programme material will be nearly maintained, since multi- generation copying is possible without deterioration. In other words a television production centre is now able to provide programmes at higher technical quality level than PAL or the other composite systems.

- The old TV transmitters constructed for black and white emission were in use for many years after the transition to colour emission, causing reduced colour quality. The new generation TV transmitters, which have gradually come into operation in the last 10 years, are very linear both in amplitude and phase. When receiving signals from these transmitters without multipath distortion, the quality is the best obtainable for a composite signal system.
- Television receivers have also been improved considerably, and are now able to give a better quality than was obtainable earlier with a composite signal. This development is due to better picture tubes and complex integrated circuitry for signal processing.
- In Norway and other mountainous countries the picture quality of the direct-to-home reception is very often considerably reduced by multipath signals. These problems can partly be solved by highly directive antennas. A totally new system is however needed to completely solve the problem. Such a system will be dealt with when discussing the coming digital TV system.

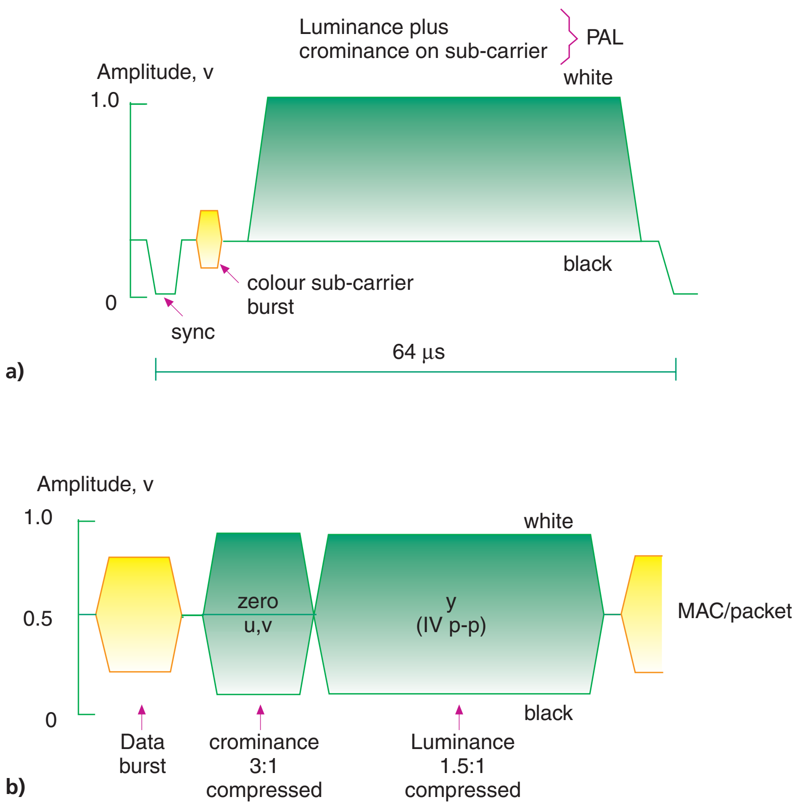


Figure 3 Line period video modulating waveform of a) PAL system, b) MAC/ packet system. Chrominance signals U and V are sent on alternate lines in MAC

3.1 Satellite broadcasting

At the frequency conference WARC 1977 channels for satellite broadcasting were allocated for countries in Europe, Africa and Asia. This plan gave rise to studies of transmission systems for improving the picture quality and the associated sound.

Around 1980 the MAC/Packet system was introduced after studies made by EBU members. MAC (Multiplexed Analogue Component) is a system where the luminance and crominance components in a picture are transmitted separately. The signal processing is digital, based on recommendation ITU-R BT.601 for digital component source coding, but the transmitted picture signal is analogue. The system has an associated data multiplex for digital transmission of the sound.

The technical quality of MAC is considerably better than PAL and the sound quality is equivalent to the NICAM digital sound in terrestrial PAL networks.

NRK started a MAC/Packet transmission by satellite in 1984 for serving the Spitzbergen Islands and the oil platforms in the North Sea with the television programme and the two sound programmes P1 and P2. Today, this transmission is carried by the broadcasting satellite TELE-X and also P3 is included in the data multiplex. NRK is using the so-called D-MAC version of the system that can carry four stereo sound channels in addition to the picture (Figure 3).

The MAC/Packet system has not been a success for satellite broadcasting in Europe for two reasons:

- The programme production up to now has mainly been based on PAL, and this composite system limits the quality of the source material.
- A large part of the population in Europe receives the satellite programme by cable operators, and all cable operators convert the MAC signal into PAL at the cable head end. This new coding into PAL gives in fact an inferior quality than would be obtained by a direct PAL satellite transmission.

The reason why NRK is using D-MAC is the ability of the system to transmit the whole NRK package of television and sound programmes.

The NRK D-MAC transmission has a beam covering the Nordic countries and is operating according to the WARC 77 plan. From the summer of 1994 a new Intelsat VII satellite will replace TELE-X for the NRK transmission, and the beam will cover large parts of Europe down to the south of Spain. Many cable operators in our neighbour countries are retransmitting the programmes. Some Norwegian cable networks and individuals are also using the satellite transmission due to serious multipath distortion by receiving terrestrial transmitters.

Most of the television transmissions via satellites in Europe are using PAL. The main reason why some programme providers have changed from PAL to MAC is that MAC can offer a more secure conditional access system, and with negligible distortion of the picture.

In conclusion, PAL will continue to be the main system for satellite broadcasting for several years to come. The replacement of PAL and MAC by a digital system will probably start in the last five years of this century.

3.1.1 Frequencies for satellite broadcasting

Two frequency bands are allocated to the broadcasting satellite service, 11.7–12.5 GHz and 21.4–22 GHz. The last frequency band was allocated at WARC 1992 for HDTV (High Definition Television).

The 12 GHz band is nearly empty, only used by few countries, because the channel plan made at WARC 77 is based on beams covering only individual countries. An exception is a Nordic beam covering the Nordic countries except Iceland. The satellite communication band 10.7–11.7 GHz is however in extensive use by programme providers feeding cable networks all over Europe. In most countries the frequency authorities have also allowed individual reception. The needs for euro-beams have in fact excluded the use of the 12 GHz band.

A replanning of the 12 GHz band is decided and will take place in a few years. A new plan will most probably give possibilities for international coverage areas, because the coming digital transmission systems will be far more frequency-effective. The new spectrum saving technique will also open the possibility of including future HDTV transmission in the 12 GHz band. The use of the 20 GHz band would be far more expensive due to the need for much higher down link transmitter power to compensate for higher atmospheric absorption and deeper signal fading.

3.2 Stereo sound in terrestrial PAL transmissions

In Germany stereo sound in television was introduced more than 10 years ago by adding a second analogue FM sound carrier in the terrestrial networks. The Nordic countries could not accept this analogue system because of the rather poor technical quality. These countries therefore started a development work in collaboration with broadcasters in England to adapt the British digital sound coding system NICAM to the B-PAL (VHF) and G-PAL (UHF) systems. In addition to the FM-modulated mono sound on a carrier 5.5 MHz above the vision carrier, the NICAM data signal is QPSK modulated on a carrier 5.85 MHz above vision carrier. This system was standardized in 1989 and is in use in all the Nordic countries and some other European countries (Figure 4).

NICAM gives in fact the TV sound quality a lead compared to FM sound broadcasting, and therefore justifies an early implementation of DAB.

3.3 New television services in the composite systems

Due to a combination of mostly historical circumstances, among which the need to have very simple receivers was probably the most important, we are using a very inefficient TV system. About 20 TV lines in each field are not in use for transmission of the picture. Some of these lines, which belong to the vertical blanking interval, are carrying test signals for checking the technical quality of the transmission. The rest is free and is today mostly used in a new data-broadcast service called Tele-text. The most widely used Tele-text system in Europe was developed by EBU members about 1980. The system can transmit pages containing 24 text lines, each consisting of 40 characters and a simple form of graphical pictures. Each page is broadcast consecutively with an instantaneous bit-rate of about 6.9 Mbit/s. The access time at the receiver depends on the amount of transmitted pages and the number of TV lines in use for the service. Today many receivers have memories for storing some Tele-text pages and thus reducing the access time for frequently used page numbers.

The present Tele-text system has rather simple characters and graphics, and must be considered as a first generation system. An EBU working group is now studying a more advanced system named Tele-text level 3, which gives much nicer

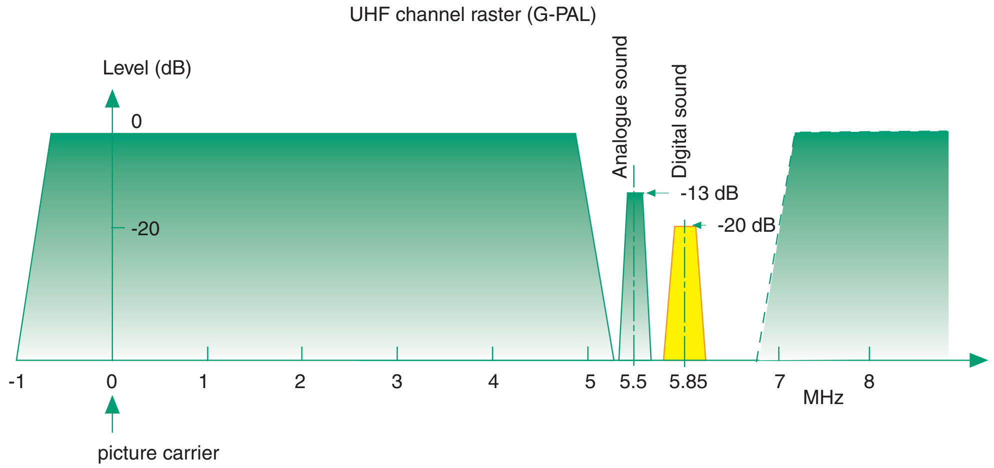


Figure 4 NICAM signal QPSK modulated on a carrier 5.85 MHz above picture carrier

characters and graphics. The system is compatible with the existing system and is therefore using the same transmission data-rate. Since each Tele-text page in level 3 contains much more data information, the access time will increase considerably. The implementation of this advanced Tele-text system is therefore dependent on more storage capacity in the receivers, but the technology is already available at reasonable cost.

NRK installed new Tele-text equipment in 1993 which is also able to handle level 3 coding, and a test transmission started in the beginning of 1994 in collaboration with the Norwegian company Teltex. Teltex has supplied the necessary decoders. Some pages of the Tele-text service is reserved for this experiment.

3.3.1 The domestic video programme delivery control system (PDC)

It is important for viewers to be able to record automatically and view TV broadcasts at a later time. Most video cassette recorders (VCR) on the market can be clock programmed for automatic start and stop according to a programme schedule announced in the newspapers. The disadvantage with this system is that the recording can start at the wrong time due to changes or delays in the scheduled time.

These facts are the background for the development by EBU members in recent years of the PDC system. PDC is based on labelling each transmitted programme giving a real time information to the VCRs for proper recording. The system is handling time delays, changes in programme schedules, interruption of transmitted programmes, time-zone boundaries, etc.

The recording-control commands may be carried in Tele-text or in a dedicated TV-line in the vertical blanking interval. The latest mentioned method is used in Germany and Austria. The other European countries are going to use Tele-text, where the PDC information is transmitted on a none visible text line (packet 8/line 30). By using Tele-text much data capacity can be saved.

Up till now very few countries have started a PDC service. The new Tele-text equipment installed at NRK is able to handle PDC signals, but the insertion of labels must be done manually at least from the beginning of the service, limited by existing installations in the main control area.

The implementation of PDC will probably take place in all the Nordic public broadcasting transmissions during 1994, except NRK, where a final decision has not yet been taken.

3.4 High definition television (HDTV)

About 20 years ago Japanese engineers believed that the time was right to add more realism to television. They performed a lot of investigations into what psycho-physical and subjectively-coloured feelings mean in terms of physical dimensions of the picture and technical parameters. They came up with a system having 1125 lines, 60 Hz field rate, line interlacing and 5:3 aspect ratio (ratio of picture width to picture height), instead of the 4:3 ratio of the existing systems.

These parameters assumed viewing on a large screen, >0.8 m^2, at a viewing distance of about three times the picture height, equivalent to a viewing angle of about 30 degrees. The normal viewing angle with existing TV systems is about 10 degrees. The wider viewing angle and wider screen give a much stronger feeling of being present. Such a TV system would match a cinema performance using 35 mm wide-screen film.

About 10 years ago a system standardization process started with the goal of obtaining a single world-wide HDTV studio production standard. It seemed, however, to be impossible to reach agreement among the countries using different field rates in existing systems, respectively 50 Hz and 60 Hz (59.94 Hz).

In Europe a system was developed based on 1250 lines, 50 Hz field rate with line interlacing, making conversion to and from the 625/50 system very easy.

The situation today is the existence of complete studio production equipment for both the 1125/60 system and the 1250/50 system. The aspect ratio is common and was chosen to be 16:9. This ratio is matching the most used wide-screen film format.

The final goal for the European system is a 1250/50 system with progressive scanning. Leaving the interlaced system will however double the system bandwidth, and in digital terms increase the bit rate from about 1 Gbit/s to 2 Gbit/s. There is, however, no doubt that a future TV system should be based on progressive scanning.

The interest among broadcasters in HDTV production is rather low for the following reasons:

- The development of a suitable transmission system has failed. In Japan a narrowband analogue system called MUSE is in operation for satellite broadcasting. But the quality is far from the HDTV studio quality, and MUSE is not in wide spread use. In Europe HD-MAC was developed, again an analogue narrow band system for satellite transmission, with a similar quality limitation as for MUSE. HD-MAC transmissions will never be implemented because the improvement in quality is not high enough. HD-MAC is based on using 16:9 receivers with conventional picture tubes, which also limit the final reception quality. 16:9 picture tubes with very high resolution are made for HDTV studio monitors, but would be far too expensive for home use.
- One of the objectives for HDTV was a large screen receiver and a close viewing distance. A conventional picture tube (CRT) would be very bulky and heavy. The only practical receiver design is a flat panel screen. The Japanese have shown a prototype of a 40 inch plasma display panel for HDTV with a thickness of 8 cm. It will, however, take at least another 5 years before mass production of such a receiver is possible.

We can therefore conclude that the home cinema, based on HDTV broadcasting with near studio quality reception, is a technique for the year 2000 or beyond.

3.5 Digital TV/HDTV

In the last part of this decade TV production based on the conventional systems (525/625 line) will be digital, and a new digital transmission standard is necessary if the improved technical quality at the studio side (EDTV – enhanced definition television) should be made available for the viewer. The term EDTV is used for 525/625 line productions based on the ITU recommendation for digital component production, Rec. ITU-R BT, 601 and 656.

This production standard has a bit rate of 270 Mbit/s, while the bit rate for HDTV is about 1 Gbit/s. It is impossible to transmit such high bit rates over satellite or terrestrial networks because of very limited available radio frequency spectrum.

Over a period of a few years video compression systems have been developed based on studies in many countries. One of the most promising activities has been the work within ISO/ MPEG. Recent results seem to indicate that subjective quality, virtually transparent to the digital component studio standard (Rec. ITU-R BT.601), can be achieved with a bit-rate of about 0.9 bit/pixel (EDTV), while 0.4 bit/pixel seems to be sufficient for a quality similar to conventional composite signals, PAL/SECAM/NTSC (defined as SDTV – standard definition television). With addition of some capacity for high quality sound, data services and error correction, a gross bit-rate in the order of 11 Mbit/s and 5.5 Mbit/s is necessary for EDTV and SDTV.

Other studies, concerning HDTV, conclude that a digital HDTV service including digital multichannel sound, data services etc., 30 Mbit/s is thought to be necessary.

Unlike many technology developments, video compression introduces fundamental changes into the market in that it permits multiple digital TV programme channels in the frequency space of one analogue programme channel, satellite or terrestrial. Or it permits a higher quality programme channel such as HDTV in the same frequency space. In summary, low bit rate transmission allows

- reduced transmitter cost per channel
- wider choice of programmes
- more channels in a given spectrum
- improved service quality including options for HDTV or multiple lower quality channels
- more secure digital conditional access systems.

In Europe the studies and standardization work of digital TV/HDTV are co-ordinated among broadcasters, telecom operators, industry, etc., and are organized according to an MOU (Memorandum of Understanding).

A basis for this standardization work is a multilayer approach. Four levels of service quality/resolution have been defined, HDTV (high definition TV), EDTV (enhanced definition TV), SDTV (standard definition TV, PAL/ SECAM) and LDTV (limited definition TV, VHS like quality). The idea is that the highest quality levels should be obtainable by stationary reception and optimal antenna installation, while the lowest quality concerns portable or mobile reception with less effective antennas, for example reception in buses, trains, etc. From an economic/ technical point of view, there is still no evidence of how many levels could be realized in the same transmission.

Another important issue in this common European study is that systems for satellite and cable transmissions should be compatible with a system for terrestrial emission in a UHF network.

The choice of channel coding and modulation is very important for an efficient use of the frequency spectrum. Satellite transmission studies have shown that a TDM (time division multiplex) approach is particularly suitable for direct to home services using small receiving antennas (60 cm). FDM (frequency division multiplex) is more interesting for multi-programme satellite distribution to cable networks head ends with larger antennas, because each FDM carrier can be distributed in the cable channels without remodulation. FDM requires more satellite transmitting power per programme channel than TDM if small receiving antennas are required. Suitable modulation systems are QPSK and 8PSK which allow the satellite power tubes to operate close to saturation, i.e. at maximum power.

With TDM and a PSK type modulation, one channel in the WARC 77 plan for satellite broadcasting in the 12 GHz band can transmit

- 1 HDTV programme
- 4 EDTV programmes
- 8 SDTV programmes.

Each country in Europe has an allocation of 5 channels in the WARC 77 plan corresponding to 5 digital HDTV programmes or 40 digital SDTV programmes.

In terrestrial and cable networks in Europe the available bandwidth per programme in PAL/SECAM is 7 or 8 MHz, corresponding to the Stockholm 1960 plan for VHF/UHF terrestrial networks.

For an HDTV service with a bit rate of about 30 Mbit/s or alternatively 4 SDTV programmes to be accommodated within an 8 MHz channel, a modulation system with a spectral efficiency of about 4 bit/s/Hz must be employed. Such a system is 16QAM, and even higher levels are under consideration (32QAM, 64QAM). These modulation systems require very high linearity of the transmitter power amplifiers, but can be realized in a practical terrestrial UHF network.

Single frequency networks (SFN), based on orthogonal frequency division multiplex (OFDM) channel coding, like DAB for radio programmes, are under study for future terrestrial UHF TV networks, in order to reduce the necessary frequency spectrum. The only available spectrum in Europe is the existing TV bands IV and V, 470 MHz – 790 MHz. These bands are, however, fully utilized in most European countries for the analogue PAL/SECAM services. When going digital, simulcast with these existing services will be necessary for many years. The only way to be able to implement digital terrestrial networks all over Europe is therefore reallocation of the frequency band 790 – 862 MHz for TV broadcasting, originally planned for television services in the Stockholm 1960 plan. After the 1979 frequency conference (WARC 79) this band has been allocated to fixed and mobile services in most European countries. The CEPT has under consideration this urgent need for frequency space for digital television.

In Norway TV 2 is running a UHF network, while NRK is mainly using VHF channels. Within a few years the NRK band I transmitters (47 – 68 MHz) will be replaced by UHF transmitters, giving space for mobile services in band I in the future. There would, however, still be enough capacity in the UHF bands to be able to implement a digital terrestrial single frequency network in Norway.

3.5.1 What does the viewer prefer?

Today it is rather obvious that the viewer would prefer a greater choice of programme channels instead of improved technical quality like HDTV. The satellite programme providers will soon offer the first preference with the introduction of the multi-programme digital television services, giving the choice of hundreds of channels. The standardization work on digital TV in Europe has given priority to satellite and cable distribution.

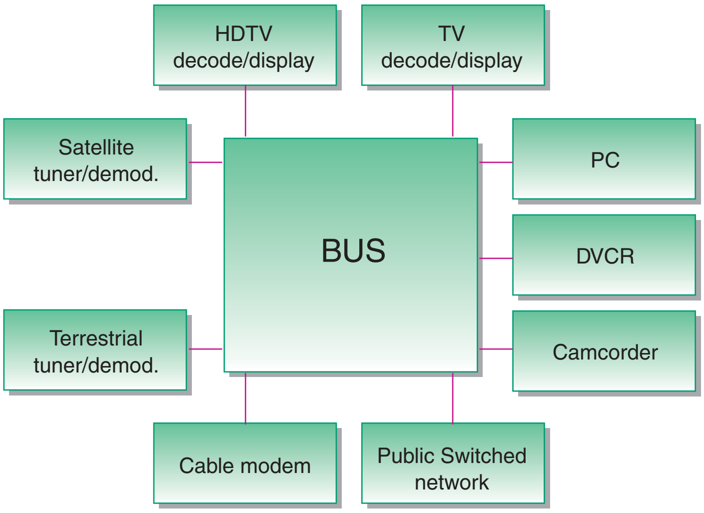


Figure 5 The multimedia home installations

Multi-channel digital satellite systems could therefore be launched already in 1995–1996. These channels will probably use the 16:9 picture format being compatible with most of the available film productions (wide screen).

In the long run a demand for more cinema like quality will come. The digital TV systems will therefore be designed for an evolution towards HDTV.

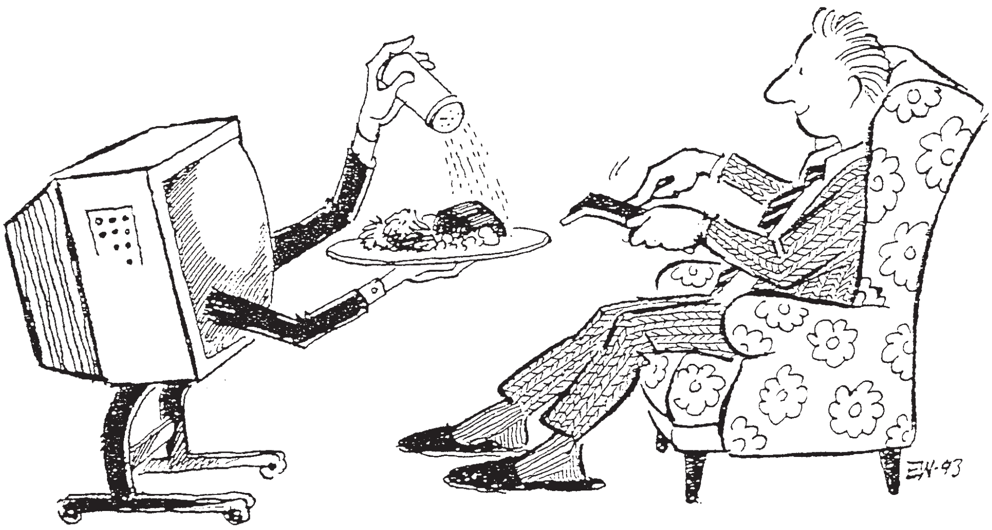
3.5.2 Is the future of television both multiple and interactive?

Hand in hand with this explosive increase in the number of programme channels, thanks to digital compression technique, the viewer will be given the means of talking back and influencing the received programmes. This is known as interactive television. An important question that has to be studied during the coming years is

- What are the possible channels to return data from the receiver to the broadcaster and other users of such data for different reception media, e.g. individual antenna, common antenna, cable, terrestrial or satellite?

A convergence of telecommunications, computer networks and TV-broadcasting is also a result of the movement from analogue to digital. The development of these multimedia systems and of interactive TV will be the great challenge in the last part of this decade. A matter of particular concern with this rapid development taking place is an orderly development of consumer equipment, such that it can cope with the added range of services likely to become available. It is expected that many homes will have several TV receivers, personal computers, VCR, video camera, video games, etc. and wish to receive and display or record programmes from them flexibly and at times simultaneously. Delivery media likely to be used include terrestrially over the air channels, cable, satellite, some delivering multiple programmes on each channel. The public switch network will most likely deliver video services as well (Figure 5).

These very interesting developments have the potential of changing the media world radically. During the next decade we will see what is going to be realized.



Perceptual audio coding for digital sound broadcasting

BY ERIK BERNATEK

System overview

The DAB transmission system for digital sound broadcasting offers robust reception of high quality sound signals for fixed, portable and mobile receivers. This article describes the sound coding technique used to reduce the amount of data that needs to be transmitted. First a short overview of the transmission system is given.

A block diagram of the transmission system is shown in Figure 1.

The DAB transmission system combines three channels:

- The *synchronization channel* is used internally within the transmission system for transmission frame synchronization, automatic frequency control, channel state estimation and transmitter identification.
- The *Fast Information Channel* (FIC) is used for rapid access of information by a receiver. In particular it is used to send Multiplex Configuration Information (MCI). The MCI describes how the multiplex is organized by defining Sub-channel organization, listing available services and managing multiplex re-configuration. Optionally Service Information and data services can be transported in the FIC which is a non-time-interleaved data channel with fixed equal error protection.
- The *Main Service Channel* (MSC) is used to carry audio and data signals. Audio and data signals are considered to be service components which can be grouped together to form services. The MSC is a time-interleaved data channel divided into a number of Sub-channels which are individually convolutionally coded, with equal or unequal error protection. Each Sub-channel may carry one or more service components.

For service components in the MSC, two different transport modes are defined:

- The *stream mode* provides a transparent transmission from source to destination at a fixed bit rate in a given Sub-channel.
- The *packet mode* is defined for the purpose of conveying several data service components into a single Sub-channel.

The main part of the MSC contains the Sound services. A sound service consists of one or more service components. The *Primary service component* carries the audio signal, and the *Secondary service components* carries additional information, e.g. Traffic Message Channel (TMC) and Service Information (SI).

The coding of Primary audio service components complies with the ISO 11172-3 standard, and is described in detail in the main part of this article.

Data services carried either in the MSC or in the FIC can be Service Information or general data services. SI features that can be carried in the FIC include language, time and country identifier, programme number, programme type, etc. The Fast Information Data Channel (FIDC) in the FIC includes provision for paging, Traffic Message Channel (TMC), Emergency Warning System (EWS), etc.

Each service component can be individually scrambled for Conditional Access (CA) to make the component incomprehensible to unauthorized users. The MCI includes information to indicate whether service components are scrambled or not, and how to find the parameters necessary for descrambling.

Each Sub-channel is individually scrambled for energy dispersal to avoid signal patterns which might result in regularity in the transmitted signal.

Convolutional coding and time interleaving is applied to the output of each energy dispersal scrambler as part of the protection against transmission errors due to adverse propagation conditions.

The convolutionally-encoded and time-interleaved Sub-channels are combined in the Main multiplexer into a structure called Common Interleave Frame (CIF). The configuration of the CIF is signalled by the MCI carried in the FIC.

The Orthogonal Frequency Division Multiplex (OFDM) modulator generates symbols from the output of the Main multiplexer. The OFDM symbols are frequency-interleaved to spread the information in each Sub-channel across the full bandwidth of the transmitted signal. The OFDM symbols are modulated by Differential Quadrature Phase Shift Keying (D-QPSK) onto each sub-carrier in the transmitted signal.

The transmission frame shown in Figure 2consists of a sequence of three groups of OFDM symbols:

- Synchronization channel symbols
- Fast Information Channel symbols, and
- Main Service Channel symbols.

The synchronization channel symbols comprise the Null symbol and the Phase reference symbol.

Audio coding

Perceptual coding techniques

Digitization of audio in the studio environment conforms to the ITU-R rec. 646 which specifies 48 kHz sampling rate and minimum 16 bit resolution. This gives a minimum bit rate of 1536 kbit/s for a stereophonic sound channel. To achieve high spectrum efficiency in the transmitted DAB signal this bit rate has to be reduced considerably. Removal of the redundancy in the audio signal (i.e. lossless data compression) is not sufficient to obtain the necessary bit-rate reduction. Irrelevancy of the audio signal (i.e. signal components that are not perceivable) is removed by exploiting properties of the human auditory system, like spectral and temporal masking.

The sensitivity of the ear varies with frequency and sound pressure level as shown in Figure 3.

Under a certain limit, called the threshold of hearing, no sound can be perceived. In the presence of a sound stimulus (e.g. narrow band noise) this threshold rises around the signal. Signal components beneath this 'masking threshold' cannot be heard. The masking threshold is dependent on the frequency and level of the masking signal (called masker) as shown in Figures 4 and 5.

When several signal components are present, the total masking threshold is the sum of the thresholds caused by the individual maskers. The character of the maskers also influences the shape and position of the masking thresholds. Tonal (sinusoidal like) and non-tonal (noise-like) maskers contribute differently to the total masking threshold, and should be treated differently in an advanced coding system. The encoding process calculates the masking thresholds and removes all signal components beneath the threshold.

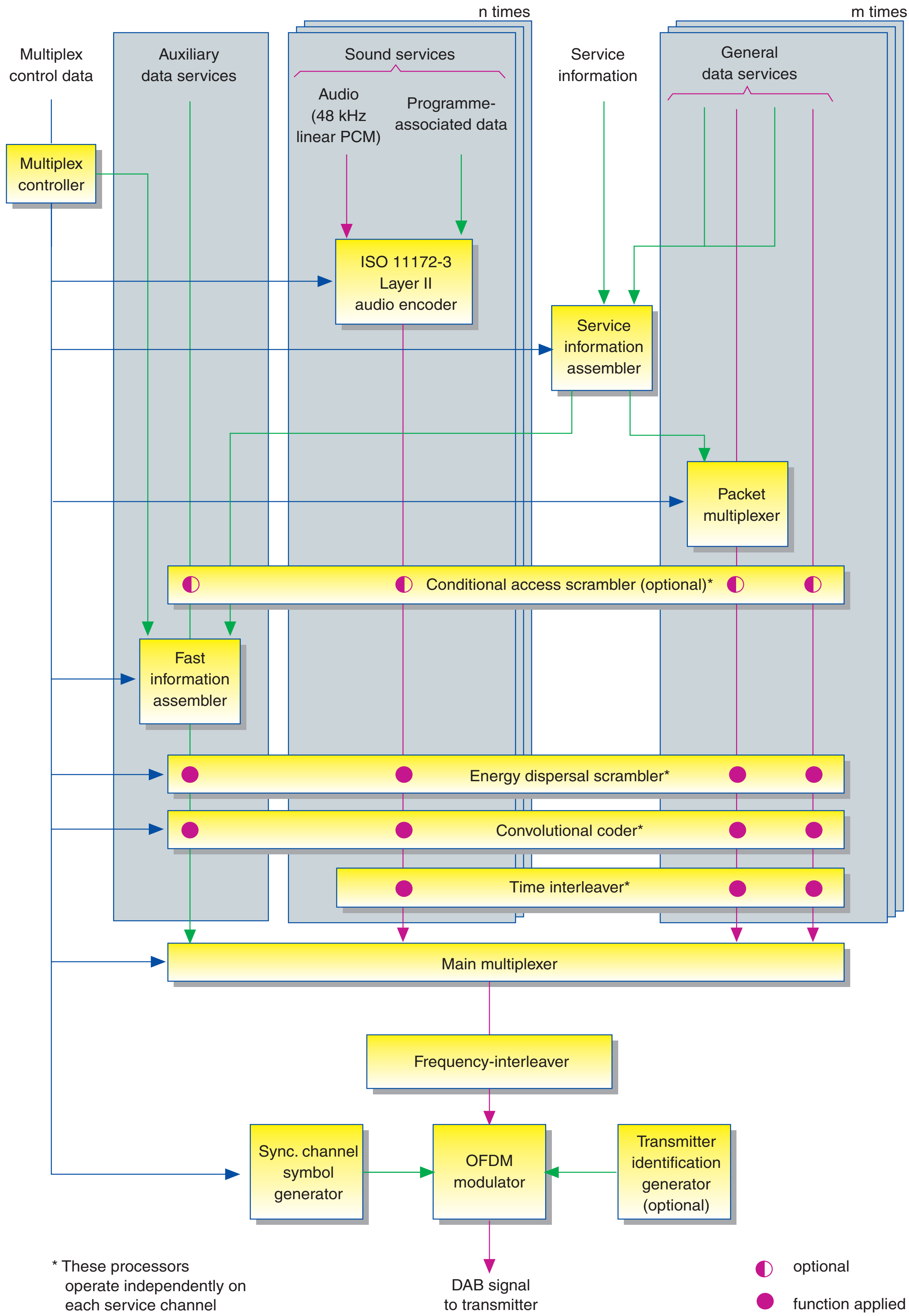


Figure 1 Conceptual diagram of the transmission part of the system

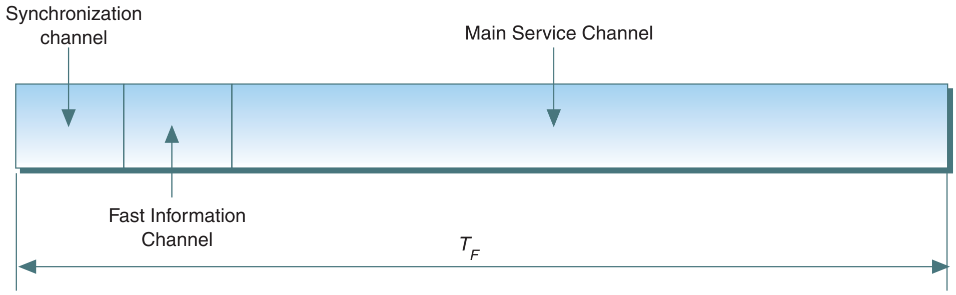


Figure 2 Description of the transmission frame

The audio samples are then quantized with the number of bits necessary such that the quantization noise will not exceed the masking threshold.

Masking in the temporal domain is shown in Figure 6. When a short stimulus is applied to the ear the sensitivity to other stimuli in a short period before (2–5 ms) and a longer period after (up to 100 ms) the main stimulus is reduced. This phenomenon can be exploited by the encoding system by removing the parts of the signal which cannot be heard due to temporal masking.

The DAB coding system

The coding of the audio services components in the DAB system complies with the ISO 11172-3 standard. The standard defines three layers of coding with increasing efficiency and complexity. Layer II is chosen for DAB. In the ISO standard only the encoded bit stream, rather than the encoder, and the decoder are specified. This makes future improvements of the encoding process possible.

Some of the main characteristics of Layer II coding are:

- Sub-band coding with 32 equally spaced sub-bands at 750 Hz
- 48 kHz input sampling frequency
- Output bit rate selectable from 32 to 192 kbit/s per monophonic channel (64 to 384 for stereo)
- No subjective degradation of the signal compared to 16 bit linear PCM at bit rates of 256 and above per stereo channel
- Four different audio modes are provided:
  - · Mono (1-channel) mode
  - · Stereo (2-channel) mode

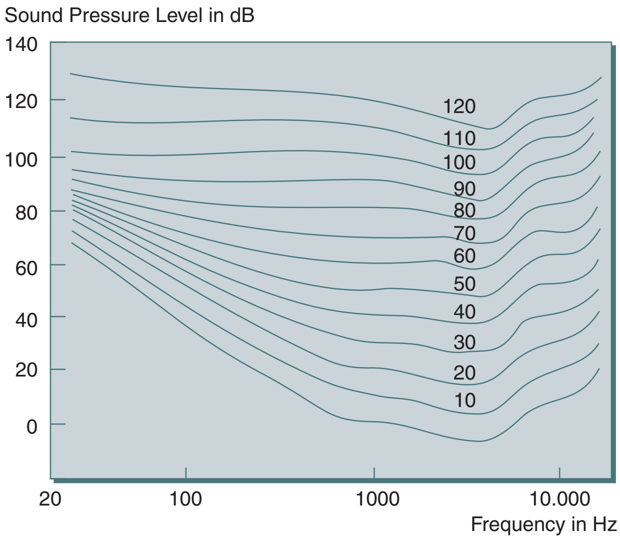


Figure 3 Fletcher/Munson loudness curves

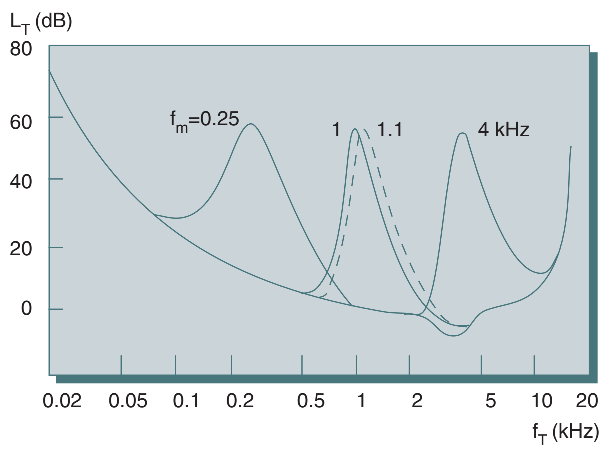


Figure 5 Masking thresholds as a function of frequency

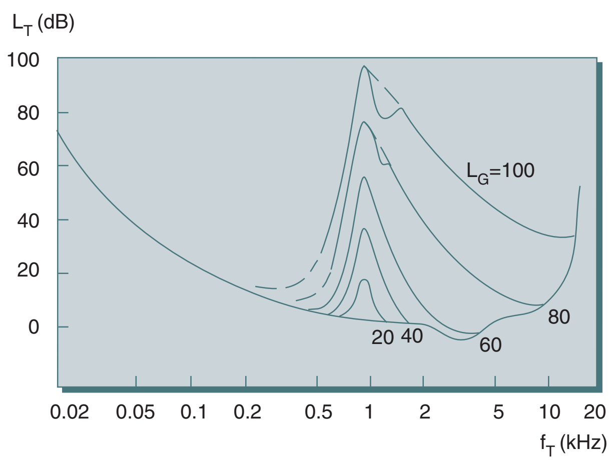


Figure 4 Masking thresholds as a function of Sound Pressure Level

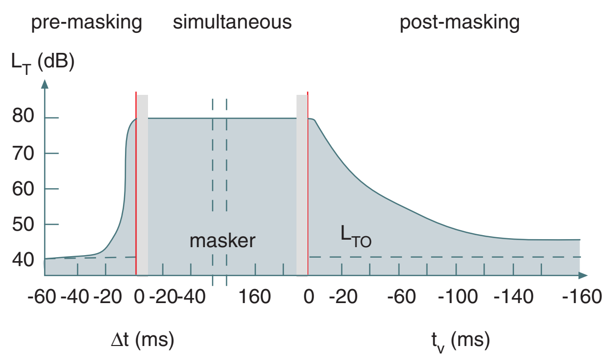


Figure 6 Temporal masking

· Dual channel mode. The two channels can be either bilingual or two mono channels

· Joint stereo mode. The redundancy and irrelevancy of stereo signals are exploited for further data reduction.

An overview of the principal functions in the audio encoder is shown in Figure 7.

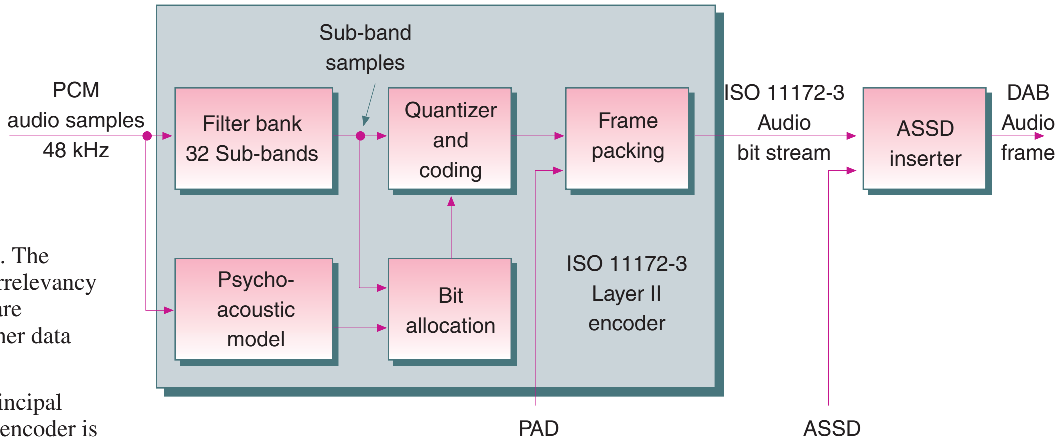


Figure 7 Simplified block diagram of the DAB audio encoder

The input audio samples are fed into the audio encoder. A polyphase filter bank of 32 sub-bands creates a filtered and sub-sampled representation of the input audio signal. The filtered samples are called sub-band samples. A perceptual model creates a set of data to control the quantizer and coding. This data can be different depending on the actual encoder implementation. One possibility is to use an estimation of the masking threshold to obtain quantizer control data. The quantizer and coding block creates a set of coding symbols from the sub-band samples. A frame packing block assembles the actual audio bit stream from the output data of the previous block, and adds other information such as header information, CRC words for error detection and Program Associated Data (PAD). Each audio frame contains all information necessary to decode the audio in the frame and has a duration of 24 ms.

The six primary parts of the encoding process are:

- Analysis sub-band filter
- Scale factor calculation
- Perceptual model
- Bit Allocation (BAl) procedure
- Quantizing and coding
- Bit stream formatter.

The primary parts of the process are shown in Figure 8 and explained in detail in the following paragraphs.

Analysis sub-band filter

The analysis filter bank splits the broadband audio signal into 32 equally spaced sub-bands. The bandwidth of each sub-band is 750 Hz at a sampling frequency of 48 kHz. This filter, called a polyphase filter, is critically sampled, i.e. there are as many samples in the sub-band domain as there are in the time domain. The filter is optimized in terms of spectral resolution with a rejection of side lobes better than 96 dB, which is necessary for a sufficient cancellation of aliasing distortions. This filter bank provides a reasonable trade-off between temporal behaviour and spectral accuracy. The requirement of having a good spectral resolution is contradictory to keeping the impulse response within certain limits. The impulse response of the filter will smear the audio signal in the time domain, and may cause pre- and post-echoes in transient rich audio material. The impulse responses must be kept short to avoid pre-echoes. Human hearing is very sensitive to pre-echoes whilst post-echoes are not as perceivable due to the temporal masking effect.

Scale factor calculation and coding

Before quantization, the output samples of the filter bank are normalized. 36 samples in each sub-band are grouped into three blocks of 12 samples each. The maximum of the absolute values of these 12 samples is determined, and coded with a word of 6 bits representing entry points to a table of quantization factors. The quantization factors cover a dynamic range of 120 dB in 2 dB steps. The audio frame corresponds to 36 sub-band samples, three scale factors are calculated for each sub-band. To reduce the bit-rate of the scale factors, the temporal masking properties of the ear are exploited. For instance, for stationary tonal sounds the deviation between the scale factors is small, and only the largest one needs to be transmitted. One, two or all of the three scale factors can be transmitted. Five classes of scale factor difference are defined and these determine the scale factor transmission patterns. The different transmission patterns are signalled in the Scale Factor Select Information (ScFSI). For sub-bands with no allocation, no scale factors are transmitted. This coding technique allows an average reduction of the bit rate of the scale factors by a factor of two.

Bit allocation and bit allocation encoding

Different strategies for allocating bits to the sub-band samples are possible. The perceptual model and the bit allocation procedure are not normative in the DAB specification. This “open” approach makes future improvements of the encoding process possible. One method, described in the DAB specification, uses minimization of the total noise-to-mask ratio (NMR) in each sub-band over the audio frame with the maximum allowed bit rate as constraint.

For an adequate calculation of the masking thresholds, a high frequency resolution in the lower frequency region, and a lower resolution in the higher frequency region is necessary. Since the polyphase filter bank divides the signal into linearly spaced sub-bands, an additional frequency analysis is needed to calculate the masking thresholds. This is obtained by using a 1024 point FFT in parallel with the sub-band filter. The output of the FFT is used to determine the relevant tonal and non-tonal maskers of the actual audio signal. The individual masking thresholds for each masker above the absolute masking threshold are calculated dependent on frequency position,

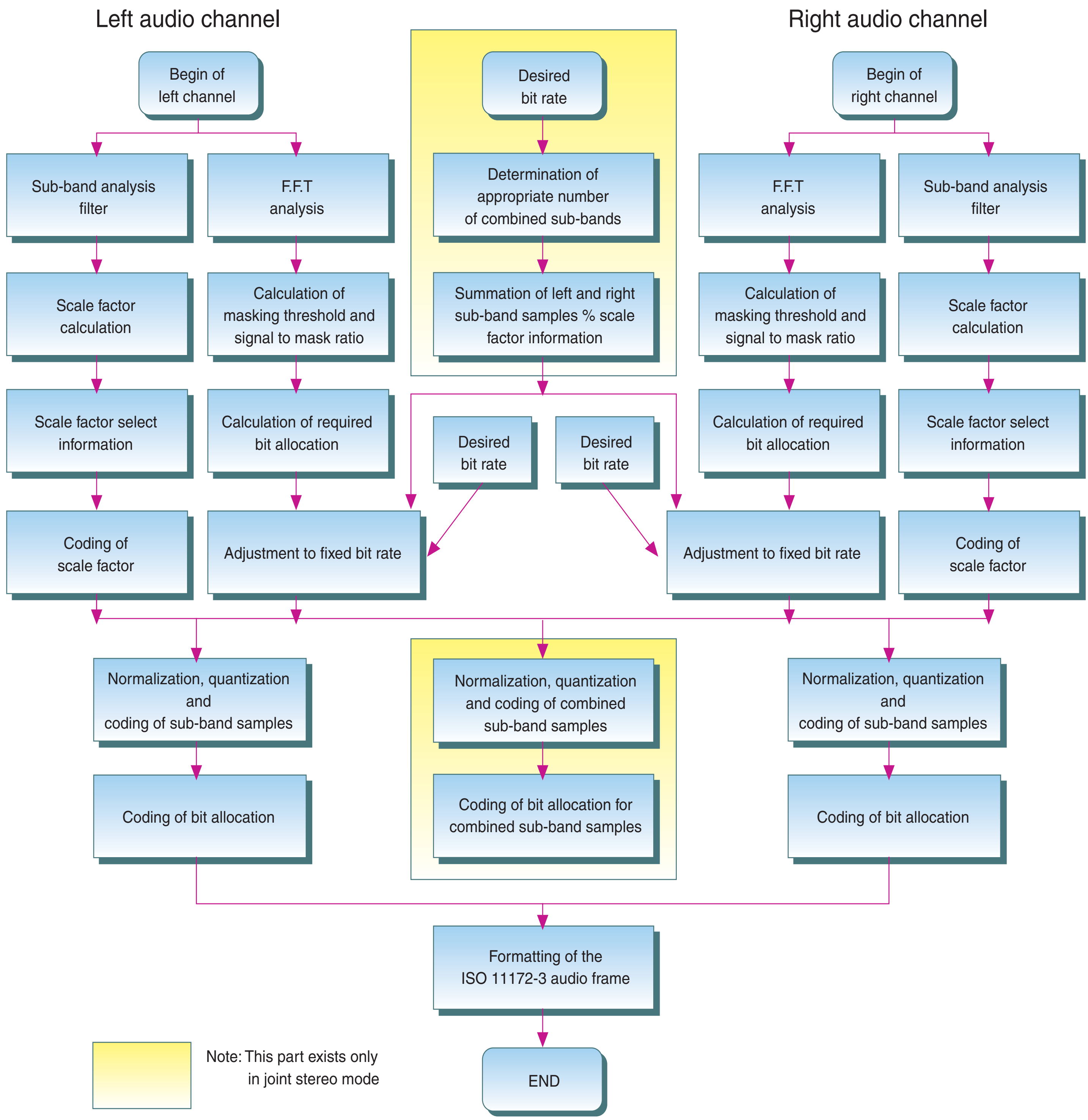


Figure 8 General ISO 11172-3 layer II stereo II encoder flow chart

loudness level and tonality. All the individual masking thresholds, including the absolute threshold, are added to the so-called global masking threshold. For each sub-band, the minimum value of this masking curve is determined. Finally, the difference between the maximum signal level, calculated from the scale factors and the power density spectrum of the FFT, and the minimum threshold, is determined for each sub-band and each block. These differences, called signal-to-mask ratios (SNR), make up the input to the bit allocation process. The allocation process is iterative. In each iteration the number of quantizing levels of the sub-band that has the greatest (perceptual) benefit is increased. The iteration is stopped when the maximum allowed bit rate is obtained. In order to increase the coding efficiency, only a

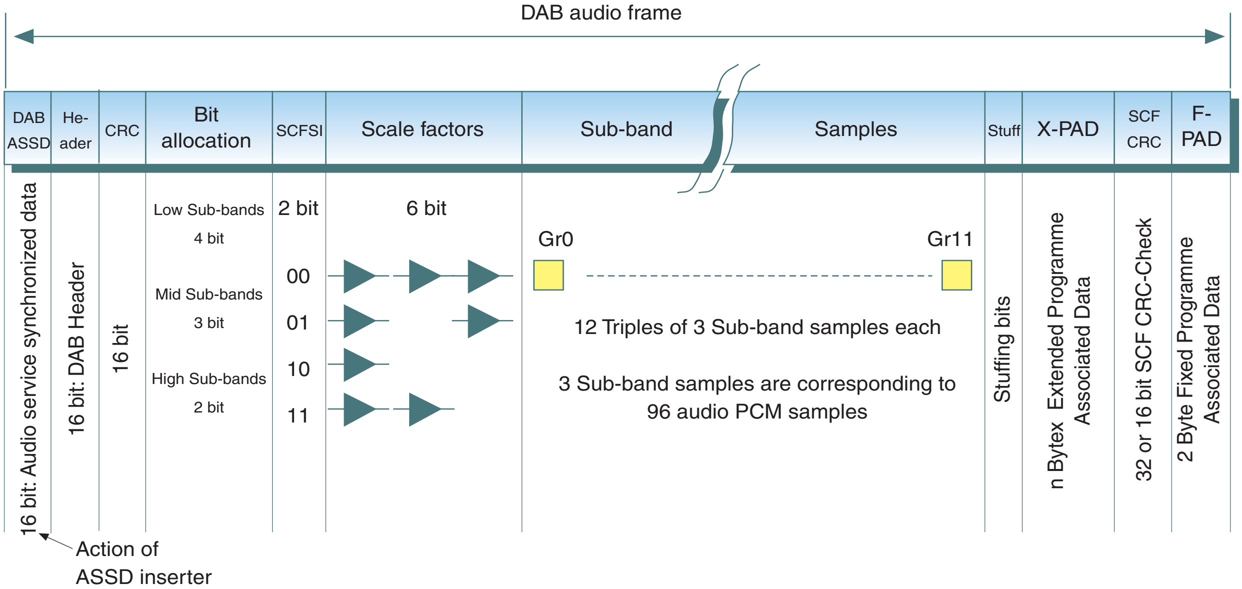


Figure 9 Structure of the DAB audio frame

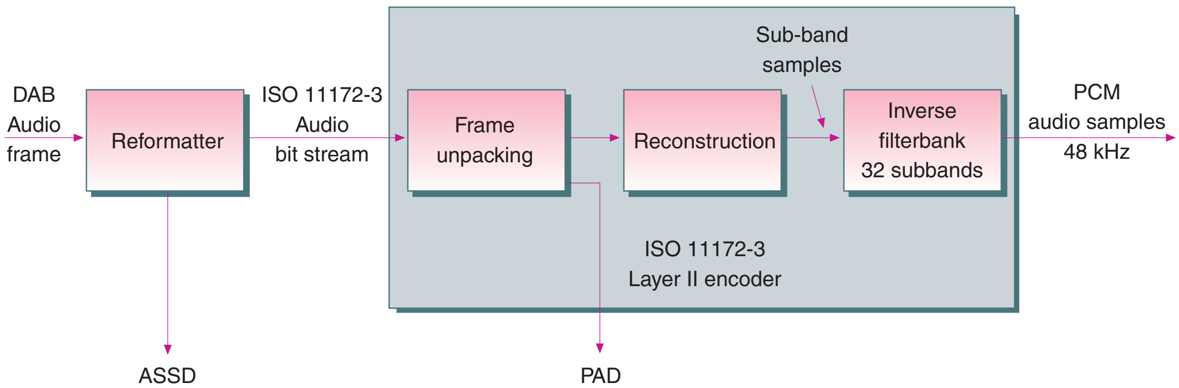


Figure 10 Simplified block diagram of the DAB audio decoder

limited number of possible quantizations are permitted. This number is dependent on the sub-band number, audio mode and total bit rate. The Bit Allocation information (BAl) is coded with two or four bits for each sub-band.

Quantizing and coding of sub-band samples

Each of the 12 consecutive sub-band samples of one block is normalized by dividing its value by the scale factor and quantized with the number of bits calculated in the bit allocation process. Only odd numbers of quantization levels are possible, allowing an exact representation of digital zero. The possible quantization levels cover the range from 3 to 65536. The number of different levels in each sub-band is however limited to 15 in the low, 7 in the mid and 3 in the high frequency bands. 3, 5 or 9 quantization level requires 2, 3 or 4 bits and will not allow an efficient use of a codeword. Therefore three successive sub-band samples, called a “granule”, are considered for coding. If all samples in one granule requires 3, 5 or 9 levels, the three samples are coded with one codeword in a process called “grouping”. In the high frequency region the samples are typically quantized with 3, 5 or 9 levels. The coding gain of using grouping is up to 37.5 %, and the overall reduction of the length of the codewords in most audio material is considerable.

The DAB audio frame structure

The structure of the audio frame is shown in Figure 9.

The bitstream is divided into audio frames, each corresponding to 1152 input audio samples, which is equivalent to a duration of 24 ms. Each audio frame contains all the information which is necessary for decoding of the bitstream. The frame starts with a header, consisting of a syncword and audio related information. A Cyclic Redundancy Check (CRC) word following the header protects a part of the header, Bit Allocation and ScFSI information. After the CRC follows Bit Allocation, ScFSI and Scale Factors. The sub-band samples are the last audio data part in the frame followed by a variable length ancillary data field. An adaption of the ISO 11172-3 Layer II is performed in order to introduce specific DAB Scale Factor Audio Check and a fixed and a variable field of Programme Associated Data (F-PAD and X-PAD).

Programme Associated Data

Each DAB audio frame contains a number of bytes carrying Programme Associated Data (PAD), i.e. information which is intimately related to the audio. All functions provided by PAD are optional. Typical examples are music/speech flags, programme related text, International Standard Recording Code (ISRC), Universal Product Code/European Article Number (UPC/EAN), a command channel which is provided synchronously to the audio programme, and a Dynamic Range Control (DRC) information. The PAD field consists of at least two bytes, called fixed PAD (F-PAD) intended to carry control information with a very strong real-time character, and data with a very low bit rate. The PAD field may be extended with a field of user defined length, X-PAD, to provide additional functions to the user, such as programme related text.

The decoding process

A simplified block diagram of the decoder is shown in Figure 10.

The decoding process requires about 1/3 of the computations needed for encoding. This is due to the fact that no perceptual model is needed to reconstruct the linear PCM audio samples. The DAB audio frame can optionally be reformatted to ISO 11172-3 format to allow use of low cost single chip decoders. Such decoders are available from several manufacturers. After reformatting, the audio frame is unpacked to recover the various elements of information. The reconstruction block reconstructs the quantized audio samples. An inverse filter bank transforms the sub-band samples back to produce linear PCM samples at 48 kHz sampling frequency. Figure 11 shows a more detailed flow chart of the decoding process.

After processing the header information, the bit allocation information is decoded in order to determine the location of the sub-band samples in the frame. The Scale Factor Select Information (ScFSI) and the scale factors are decoded before requantization of the sub-band samples. 32 sub-band samples, one from each sub-band, are fed into the synthesis filter bank which outputs 32 linear PCM samples. Since the filter produces samples that are consecutive in time, the overall delay in the coding process is mainly determined by the length of the impulse responses of the polyphase filters. The filter structure is extremely efficient for

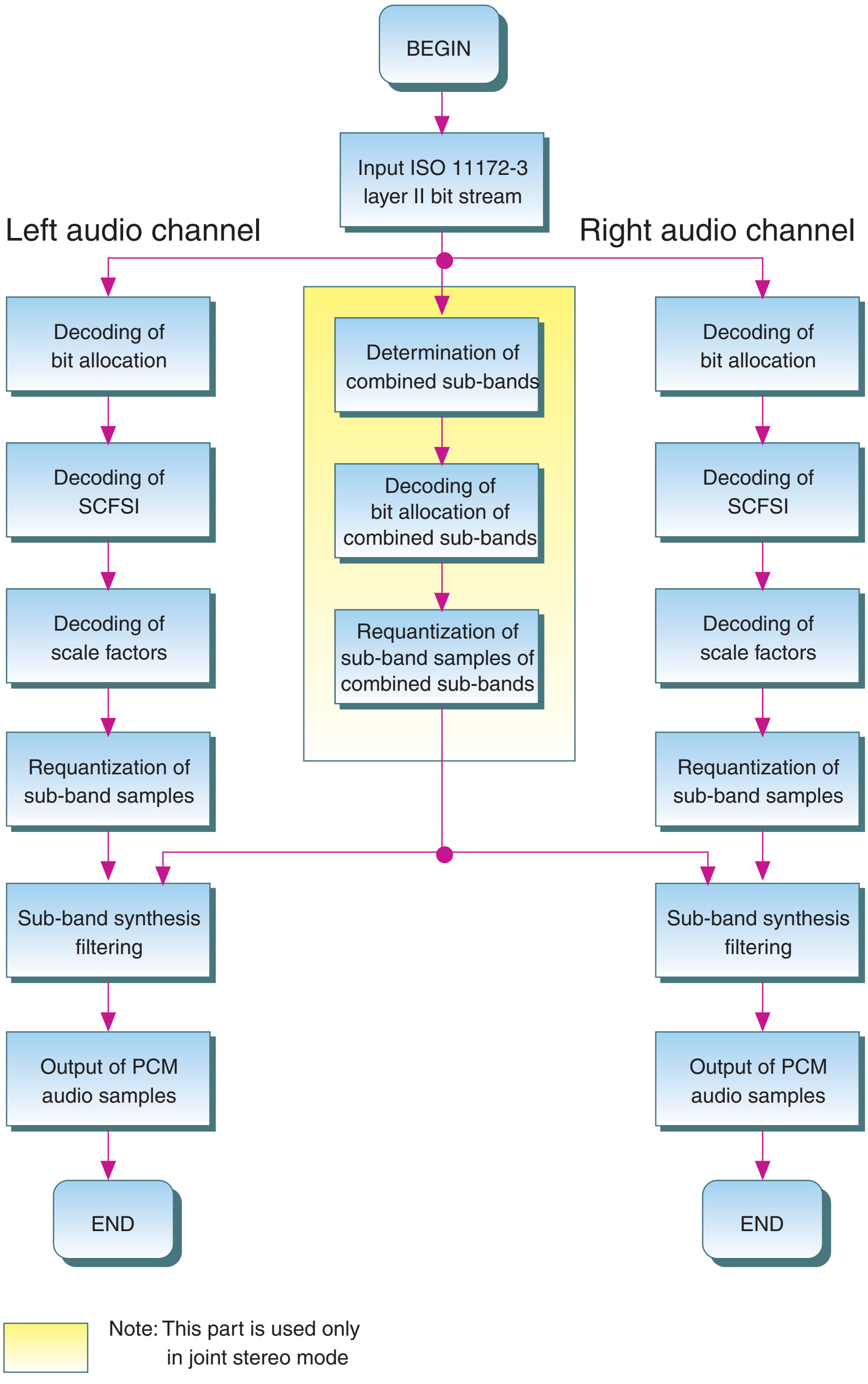


Figure 11 General ISO 11172-3 layer II stereo II decoder flow chart

implementing in a low complexity and non-DSP based decoder and requires generally less than 80 integer multiplications/additions per output sample. Moreover, the complete analysis and synthesis filter gives an overall time delay of only 10.5 ms at 48 kHz sampling rate.

Digital Sound Broadcasting – Transmission Systems and Terrestrial Networks

BY EIVIND VÅGSLID SKJÆVELAND

1 Introduction

The FM radio has been the dominating system for sound broadcasting for the last 40 years, and it can provide good sound quality for stationary reception. Despite recent improvements, as FM stereo and RDS, the FM system has its shortcomings. FM reception in a car is difficult. Noise, interference and reflected signals will often degrade the signal. New sound media, like the CD record and NICAM sound in television, provides a significantly better sound quality than FM.

In most European countries there is little space left for additional programmes in the FM band 87.5 – 108 MHz. It is not feasible to expect that any more frequency space will become available for broadcasting services. The only way to increase the number of programmes is on the cost of increased interference, and thus decreased sound quality.

If a new system is to replace the FM radio, it has to offer a sound quality at least as good as the other new sound media. It must also be well suited for mobile reception. To be able to provide more programmes, the spectral efficiency must be significantly better than in the FM radio.

Certain digital transmission techniques can give this improved spectral efficiency. In addition, digital transmission can provide better ruggedness against noise, interference and reflected signals, which is a particular problem for mobile reception. It will also be possible to improve the sound quality to the same level that we experience with e.g. CD records.

The European co-operation project Eureka 147 has developed a system called Digital Audio Broadcasting (DAB®). It seems to have the properties one must expect from a system that aims at replacing the FM radio.

2 Mobile and Stationary Reception

2.1 Multipath transmission

A common situation for radio reception is where the signal takes several paths from a transmitter to the receiver. This is called multipath propagation, and it usually leads to degradation of the received signal. Besides the direct path from the transmitter to the receiver, there are several indirect paths, caused by signals reflected from buildings, mountainsides, vehicles, etc. Often, there is no direct path present, only reflected signals.

During stationary reception, one can usually overcome this problem by using a directive antenna, which enhances the signal strength in the direction of the strongest signal, and suppresses signals from other directions. In a car, however, one cannot use a directive antenna, and the problem of multipath is more severe.

The motion of the mobile receiver will cause the received signal to be Doppler shifted in frequency. (1) It will be shifted upwards in frequency if the car moves towards the apparent signal source, downwards if the car moves away from the source. In a multipath situation, the different paths will give different Doppler shifts to the signal. The receiver will receive copies of the same signal at slightly different frequencies. This may cause degradation of the signal.

The main effect of the multipath channel is *fading* and inter-symbol interference. The signal strength of the received signal can be attenuated, and may occasionally disappear totally. Attenuation of the signal will vary strongly within even a few metres. A mobile receiver will experience this as a time-varying fading of the signal.

The multipath channel will introduce inter-symbol interference on a digital signal, due to the spread of the time delay in the various signal paths. Figure 2 shows an example of how the amplitude response may vary as a function of both time and frequency. A transmission system used for broadcasting to mobile receivers has to cope with such channels.

The fading in these channels may be frequency selective, which means that the fading of the signal components varies with frequency within the bandwidth of

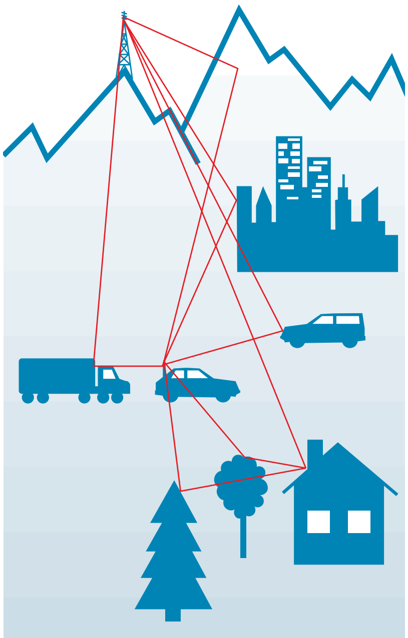


Figure 1 Example of a multipath channel. The radio signals reach the receiver (in the car) via multiple reflections

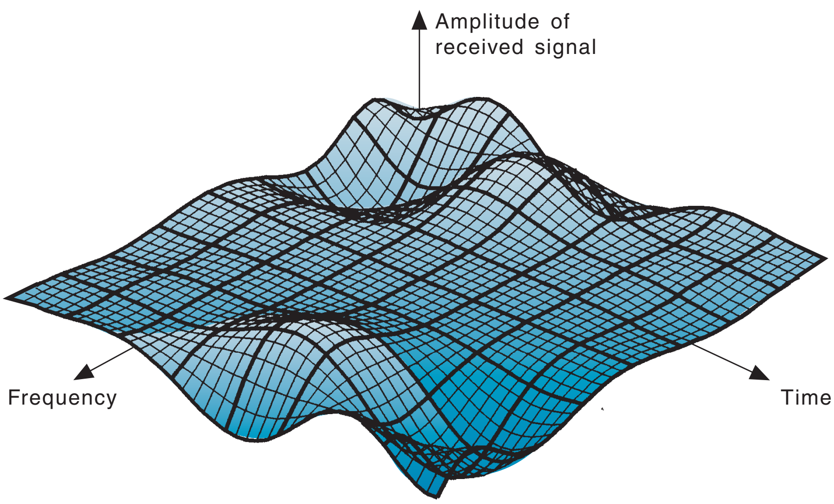


Figure 2 Amplitude response of a typical time-varying mobile channel

the channel. This selectiveness introduces distortion and inter-symbol interference to the signal. In a non-selective or flat fading channel, the fading is constant within the channel, and will only give an attenuation to the signal.

2.2 Modulation for a multipath channel

Narrowband channels, which carries a low bitrate signal, are less vulnerable to inter-symbol interference. Figure 3 illustrates this effect. On the top we see the impulse response of the channel. It shows that the delay of the different paths of the signal is spread over a time τ. Furthermore, we see the symbol trains of two different digital signals. The high bitrate signal has a symbol time t_{sh}, and a low bitrate signal with symbol time t_{sl}. $t_{sh} < \tau$ while $t_{sl} >> \tau$. For the high bitrate signal, each symbol will interfere with the two following symbols. For the low bitrate signal, each symbol will only interfere with a small fraction of the next symbol. Inter-symbol interference is thus a much smaller problem for the low bitrate signal.

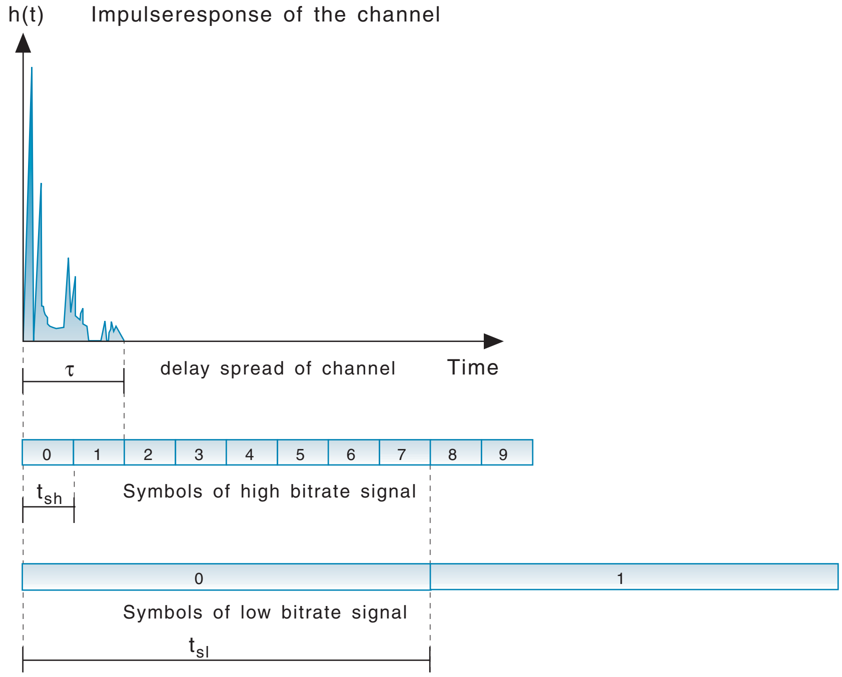


Figure 3 Impulse response of a multipath channel

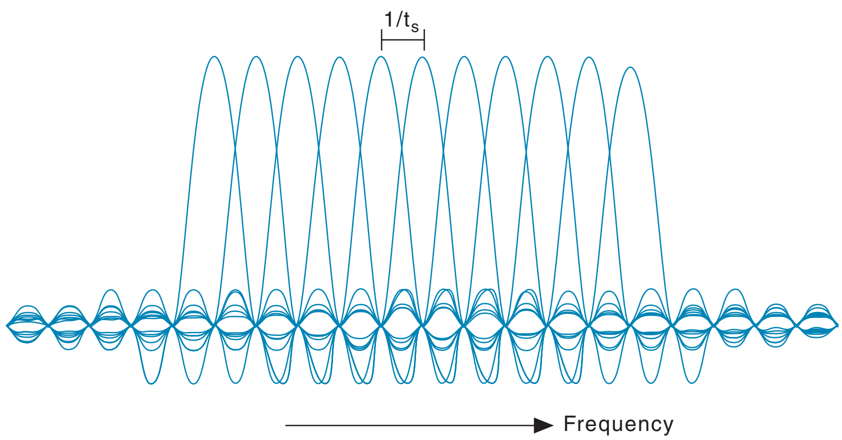


Figure 4 Frequency domain representation of orthogonal multiplexed QPSK carriers

Stretching this example a little further, we see an approach to overcome the problem of selective fading. (2) The information in the high bitrate signal is split into many low bitrate signals. Each low bitrate signal is modulated onto a carrier in a narrowband channel. The carriers are very closely spaced in frequency, together they fill a large bandwidth. The frequency selective wideband channel is transformed into many sub-channels, and each sub-channel is non-selective. In figure 2 the small squares represent the sub-channels. Within each square, the amplitude response is approximately uniform.

"Orthogonal Frequency Division Multiplexing" (OFDM) is a particularly efficient way to modulate and multiplex the many sub-channels. In each sub-channel we have a QPSK modulated carrier. The symbol time of the QPSK signal is t_s. Each carrier has a frequency spectrum with zero-crossings at integer multiples of $1/t_s$ from the carrier's centre frequency. By placing the carriers of the other sub-channels exactly at the zero-crossings, we have an orthogonal system.

In figure 4 we can see a number of QPSK carriers multiplexed this way. The carriers are spaced in frequency by $1/t_s$. At the point of the centre frequency of one carrier, all the other carriers have a zero amplitude. Thus, each carrier can be demodulated without interference from the other carriers.

The orthogonality removes the need for a guard band between the sub-channels. We can stack carriers together very efficiently, and achieve a very good use of the frequency spectrum. When modulating each carrier using QPSK, we get a spectral efficiency of 2 bit/s/Hz.

Because of the low symbol rate in each sub-channel, the symbol time can be long. As we can see in figure 3, the problem of inter-symbol interference is reduced, but is not totally eliminated. In OFDM we solve the problem of inter-symbol interference by introducing a *guard interval* between each symbol. Figure 5 illustrates this. If the duration of the guard interval (Δ) is longer than the duration of the impulse response of the channel (τ), all reflected copies of the previous symbol will be received before the end of the guard interval. Thus, we have no inter-symbol interference. The reflections of a symbol will only interfere with other reflections of the *same symbol,* and the interference will be mainly *constructive.* The cost of using a guard interval will be that of less efficient information transfer, since the guard interval carries no useful information.

Use of the OFDM technique and the guard interval take care of inter-symbol interference, but we still have to solve the problem of flat fading in each sub-channel. This will lead to attenuation of the signal in the sub-channel, and in some sub-channels, to transmission errors and loss of information.

Fading is generally selective with respect to the full bandwidth of the OFDM channel. This means that the fading will affect each sub-channel differently. When the frequency distance between two sub-channels is sufficiently large, the fading in the two channels can be considered independent of each other. Since the mobile channel is time variant, the fading in one sub-channel is also independent of fading in the same channel at another time. This is illustrated in figure 2. The large squares show the distance, in time and frequency, between points in the time frequency domain considered statistically independent.

This property can be used to recover lost information, by using a powerful error correcting code, combined with interleaving of the information in both time and frequency.

A convolutional code adds redundant protection bits to the information bit stream. After the convolutional coding, we have a bit stream where nearby bits are correlated. This redundancy and correlation allow the receiver to correct transmission errors. Interleaving in time and frequency means that sequential bits are transmitted separated in time and at different sub-channels. It becomes more likely that correlated bits are put in statistically independent sub-channels and symbol intervals. Thus, it is unlikely that a burst of errors, corrupting several bits, will affect correlated bits. Our chances of correcting the errors at the receiver are better.

Referring back to figure 2: We assume that each of the small squares are symbol intervals in a sub-channel. Within this square the channel is assumed invariant. The large squares show the minimum distance between sub-channels and symbols that are assumed statistically independent. The interleaving makes sure that all bits transmitted within the same large square are uncorrelated. Correlated bits are transmitted in different large squares. A deep fade resulting in bit errors will probably affect the symbols in one or a few large squares. The protection bits we need to correct the errors, will be transmitted in other large squares, and are probably unaffected by this fade. In most such cases, it will be possible to correct the errors.

Coding and interleaving are essential to the OFDM system to work properly. The coding is therefore often regarded as an integrated part of the system, called “Coded Orthogonal Frequency Division Multiplex”, or COFDM.

Figure 6 illustrates the COFDM signal in both the time and frequency domain. It shows how sequential bits are relocated in time and frequency after interleaving. (3)

Important parameters of the COFDM signal are: symbol time, duration of the guard interval, frequency spacing

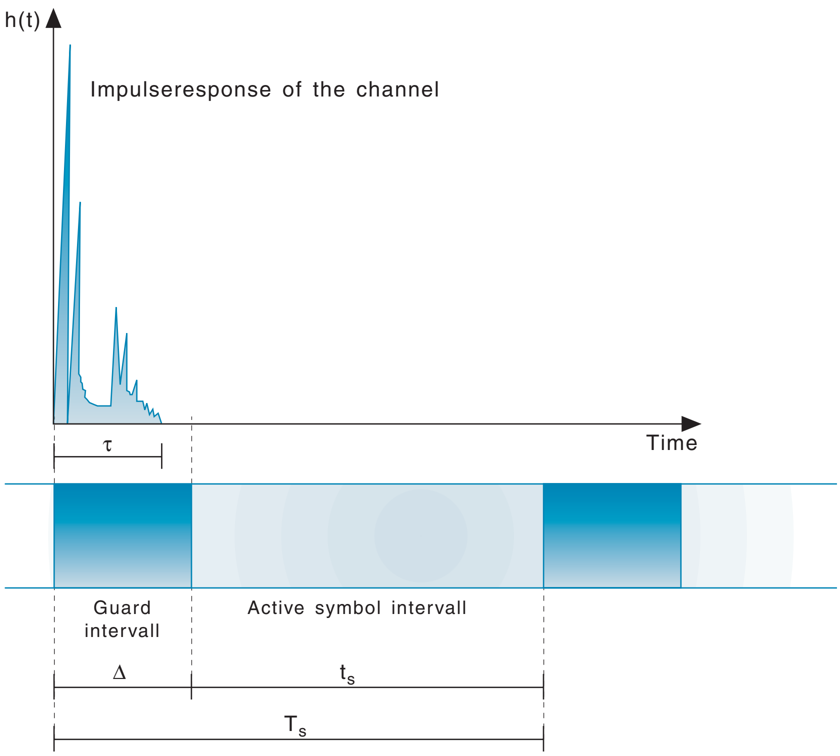


Figure 5 The guard interval

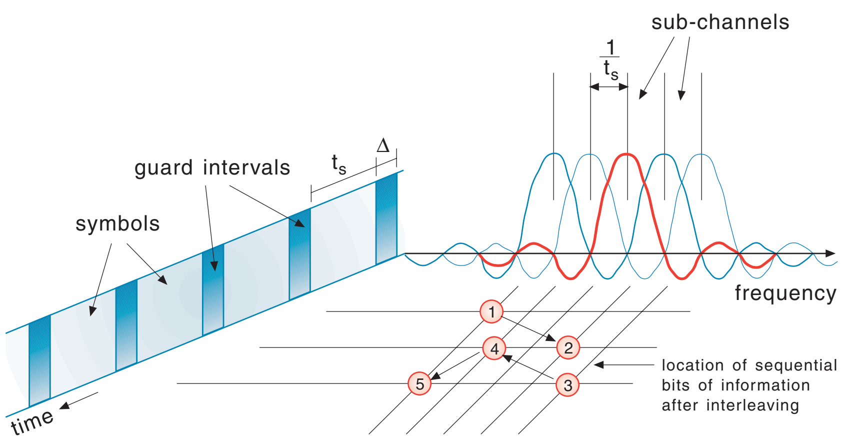


Figure 6 Spectral and temporal representation of a COFDM signal (3)

between the carriers of each sub-channel, and the total bandwidth of the signal.

To be sure that the sub-channels are frequency invariant, (4) the active symbol time t_s should be about 10 times longer than the delay spread of the cannel. The guard interval (Δ) should exceed the expected duration of the channel's impulse response. To avoid reducing the transfer efficiency too much, the guard interval should not exceed 20 % of the active symbol time t_s.

Doppler shift sets the limit of how narrow the spacing of the carriers can be. The maximum Doppler shift must be significantly less than the carrier spacing (e.g. less than 1/25th of the carrier spacing). As the carrier spacing is the inverse of the active symbol time t_s, the Doppler shift sets the upper limit on the symbol duration.

Doppler shift is frequency dependent, increasing with increasing transmitting frequency. On a high transmitting frequency, the carrier spacing needs to be larger, and thus the symbol time must be shorter than it can be on a lower frequency. A long delay spread requires a long symbol time, and thus the delay spread of the channel limits the transmitting frequencies we can use.

The frequency interleaving requires a rather large bandwidth, more than 1 MHz, to be efficient. This is more than we need to transmit only one sound programme. To preserve the frequency efficiency, we need to multiplex several sound programmes together in one COFDM signal.

3 The Single Frequency Network

(5) The COFDM technique we have described in the previous chapter, is well suited in environments in which one must expect many reflected signals. In many cases, it will not only tolerate the reflections, but even benefit from them. Due to the guard interval, the reflected signals give constructive interference, and improve the quality of the received signal. This property opens new possibilities for planning of transmitter networks.

The receiver sees no difference between a passive echo caused by reflections, and an active echo, generated by another transmitter. In both cases it will receive several delayed copies of the transmitted signal. If the differences in delay do not exceed the length of the guard interval, and the signals are identical and synchronous, the echoes will add up constructively.

This allows us to use the same frequencies for several transmitters in an area, given that they transmit the same signal synchronously. This network, called a *Single Frequency Network* (SFN) is very frequency efficient. For national or subnational coverage, the SFN can transmit from 4 to 12 times more programmes than a traditional FM network, within the same bandwidth.

The SFN demands that every transmitter in the network transmits exactly the same COFDM signal at exactly the same time and at exactly the same frequency. Thus, the network cannot be divided into (regional or local) sub-networks for parts of the time. We need separate local, regional and national networks, or regional networks that also transmit national programmes. The SFN is less flexible than the current FM networks, in the case of geographical configuration. This is the cost of increased spectral efficiency.

The maximum difference in signal delay is decided by the spacing of the transmitters in the network, and the transmission parameters of the COFDM signal must be selected according to this. If the transmitter spacing is 60 km, the delay spread will be approximately 200 µs, and the guard interval must exceed this.

4 The Eureka 147 DAB® system

This chapter describes the main features of the Eureka 147 proposal to a digital sound broadcasting system, the DAB® system. The system has been proposed as a European Telecommunication Standard (ETS) (6) for digital sound broadcasting.

4.1 Transmission system

The DAB system uses COFDM modulation with QPSK modulation of each carrier, the modulated signal occupies a nominal bandwidth of 1536 kHz.

The Doppler shift the receiver is supposed to tolerate sets an upper limit on the symbol time. As the Doppler shift is frequency dependent, the upper limit of the symbol time is also frequency dependent. Due to this, the Eureka 147 has defined three different transmission modes of the DAB signal, each mode having transmission parameters optimized for particular operating frequencies and network configurations.

Table 1 shows the main transmission parameters of the DAB system in the three different modes.

Mode I is intended for use in networks operating in the VHF frequency range, 50–250 MHz, and is particularly suited for single frequency networks. Mode II is intended for use for local broadcasting in the VHF and UHF bands and satellite broadcasting at frequencies below 1.5 GHz. Mode III is intended for use in satellite and terrestrial broadcasting at frequencies below 3 GHz.

4.2 Audio coding

The Eureka 147 DAB system uses ISO 11172-3 layer II audio coding. (7) This coding technique uses an algorithm that takes advantage of certain properties of the human sound perception. It allows a bitrate reduction from 768 kbit/s down to about 100 kbit/s per mono channel while the subjective quality of the sound is preserved. The bitrate can take on a number of values ranging from 32 kbit/s to 192 kbit/s per mono channel, or 64 to 384 kbit/s per stereo program. The lowest bitrates are achieved at the expense of reduced sound quality.

Table 1 Transmission parameter of the Eureka 147 DAB system

	Mode I	Mode II	Mode III
Nominal Bandwidth	1536 kHz	1536 kHz	1536 kHz
Number of transmitted carriers	1536	384	192
Carrier spacing	1 kHz	4 kHz	8 kHz
Total symbol time TS	1246 ms	312 ms	156 ms
Guard interval	246 ms	62 ms	31 ms

4.3 Multiplexing

The modulated COFDM signal of the DAB system has a gross transmission capacity of 2.3 Mbit/s. Several sound programmes share this capacity in a time division multiplex configuration.

The information of each programme is protected by a convolutional code. Coding introduces redundant bits in the audio bit stream. The code rate is defined as the ratio of information bits to total number of bits after coding, and the level of protection is dependent on the code rate. A low code rate gives a good protection against transmission errors, at the expense of transmission capacity available for information bits.

The number of sound programmes that can share the same multiplex is depending on both the audio bitrates and the code rates selected for each programme. We have a trade-off between the number of sound programmes in a DAB multiplex on one side, and sound quality and error protection on the other.

By using an audio bitrate of 192 kbit/s per stereo programme and a coding rate of 1/2, there will be room for six programmes in a DAB multiplex.

5 A possible network solution

In this chapter we propose a structure for a DAB network in Norway, using the Single Frequency Network concept. (5) The aim is to find a network structure that can offer both regional programmes for each county (fylke) and national programmes.

This proposal is based on some assumptions (that might not be correct):

Audio coding
It is assumed that an audio bitrate of 192 kbit/s per stereo programme gives a sufficient audio quality. Using this bitrate, each DAB multiplex can contain six stereo programmes.

Available frequencies
It is further assumed that sufficient frequency space can be made available. The proposed network needs four frequency blocks of 1.75 MHz each, including guard bands between the blocks, a total of 7 MHz. This bandwidth should be in the VHF range (30–300 MHz), as this is the optimum frequency range for SFN. (6)

Frequency planning and co-ordination
The planning of the network has to be co-ordinated with neighbouring countries, to avoid interference across the borders. The limitations on the network structure this co-ordination will impose, is not considered here.

An important requirement of the SFN is that all transmitters have to transmit the same signal simultaneously on the same frequency. The transmitters of the SFN transmit a time-division multiplex of six programmes. It has to be the same six programmes for all the networks. This makes sharing, even time sharing, between national and regional networks, impossible. A national SFN cannot transmit different regional programmes in different regions. The network structure ought to be a mix of both permanent national networks and permanent regional networks, in order to have both national and regional programmes. We can, of course, have only regional SFNs that transmit both regional and national programmes, but this is a less spectral efficient way to transmit the national programmes.

The network proposal will offer two regional programmes per county and six national programmes. A national SFN, using one frequency block all over the country, transmits one DAB multiplex with the six national programmes. The regional programmes are transmitted over seven regional SFNs. Each regional SFN covers two or three counties. As the DAB multiplex in each SFN transmits six programmes, we get two or three programmes per county.

The transmitted signal has to be the same in the whole SFN area. The programmes transmitted in a regional SFN covering several counties can be received in all of the covered area, even if the programme content is aimed at a single county.

By reusing the frequencies in regions far apart, the seven regional SFNs need only three frequency blocks. This is illustrated in figure 7. The map shows the regions, numbered from 1 to 7. The frequency blocks are named A, B and C, and it is indicated on the map which frequency block the SFN in each region is using. In table 2 the regional networks are summarized.

This network gives eight programmes (six national and two regional) in any county. The necessary bandwidth is 7 MHz, or about 900 kHz per programme. To achieve the same number of programmes in an FM network, we would need between 25 and 30 MHz.

We have not considered local broadcasting for city or town coverage, in this proposal. The frequencies used for the regional SFN in one region can, in some places, be reused for local broadcasting on low power transmitters in the other regions. But this is insufficient in order to offer local DAB broadcasting in every community. To achieve this, more frequencies are needed.

Table 2 Regional Single Frequency Networks

Region	Frequency block	Counties
1	A	Finnmark, Troms, Nordland
2	B	Nord-Trøndelag, Sør-Trøndelag, Møre og Romsdal
3	A	Sogn og Fjordane, Hordaland
4	C	Rogaland, Vest-Agder, Aust-Agder
5	C	Hedmark, Oppland
6	B	Telemark, Buskerud, Vestfold
7	A	Østfold, Akershus, Oslo

6 References

1 Jakes, W C. *Microwave mobile communications.* New York, John Wiley, 1974.

2 Alard, M, Lassalle, R. Principles of modulation and channel coding for digital broadcasting for mobile receivers. *EBU Review,* 224, 1987.

3 Maddocks, M C D. *An introduction to digital modulation and OFDM techniques.* Kingswood Warren, England, The British Broadcasting Corporation, Research Department, 1992.

4 Le Floch, B, Halbert-Lassalle, R, Castelain, D. Digital Sound Broadcasting to mobile receivers. *IEEE Transactions on Consumer Electronics,* 35, 493–503, 1989.

5 Kuhl, K H, Petke, G, Schneeberger, G. *Zur Gestaltung von Sendernetzen für eien digitalen terrestrischen Hörfunk.* München, Institut für Rundfunk technik, 1990.

6 DRAFT prETS 300 402. *Radio broadcast systems: digital audio broadcasting (DAB) to mobile, portable and fixed receivers*. Sophia Antipolis, EBU/ETSI JTC, januar 1994.

7 International Standard ISO/IEC 11172-313. *Coding of moving pictures and associated audio for digital storage media at up to 1.5 Mbit/s – audio part.* March 1993.

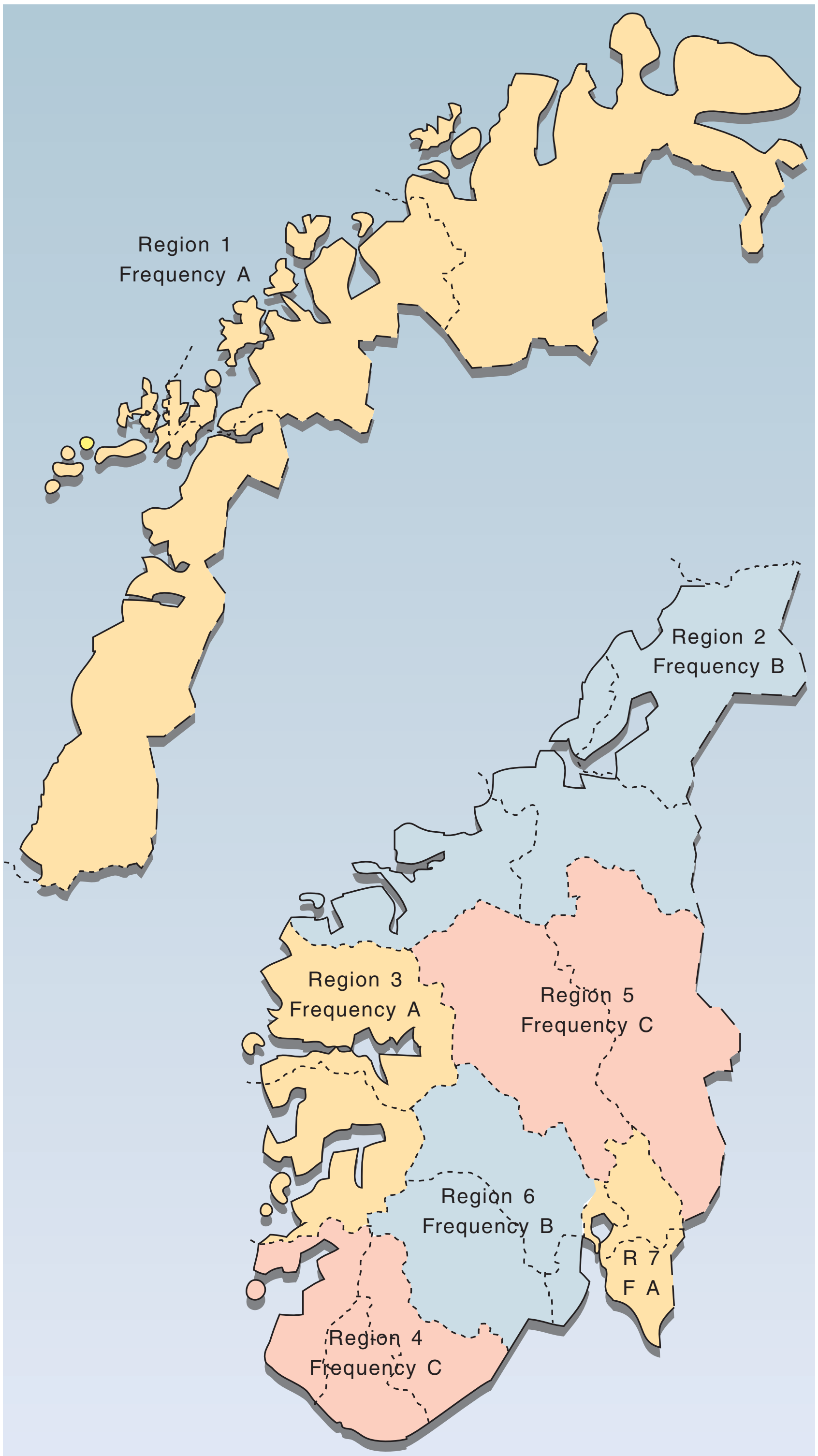


Figure 7 A proposed structure of regional Single Frequency Networks

Digital Audio Broadcasting: Terrestrial transmission experiments in Norway

BY ROLF MØKLEBY

Digital Audio Broadcasting (DAB) is a new system with possibilities and restrictions different from other radio broadcasting systems. Many experiments have been carried out in other countries in Europe as well as in Canada. Most of these experiments have been concerned with testing the theory of the system. Many of the tests have been carried out with low power transmitters, often covering urban areas.

In Norway there are long distances between main transmitters, and several smaller fill-in transmitters (repeaters) are needed to cover the fiords and valleys. In such areas the signal is normally disturbed by reflections and the level has a wide range in any given area. It is therefore important for us to test the system with realistic power levels.

The objectives for the DAB test network in Norway are:

a To install a complete DAB chain with encoding and transmission in a Single Frequency Network (SFN), to gain experience with DAB signal handling and the associated equipment

b To study coverage for mobile and stationary reception under difficult topographical conditions for both horizontal and vertical polarization of transmitted signals

c To perform demonstrations.

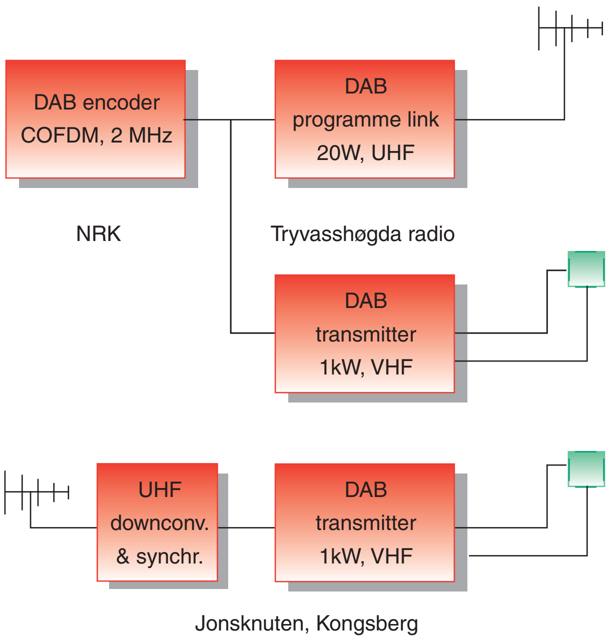


Figure 1 Test network DAB Norway

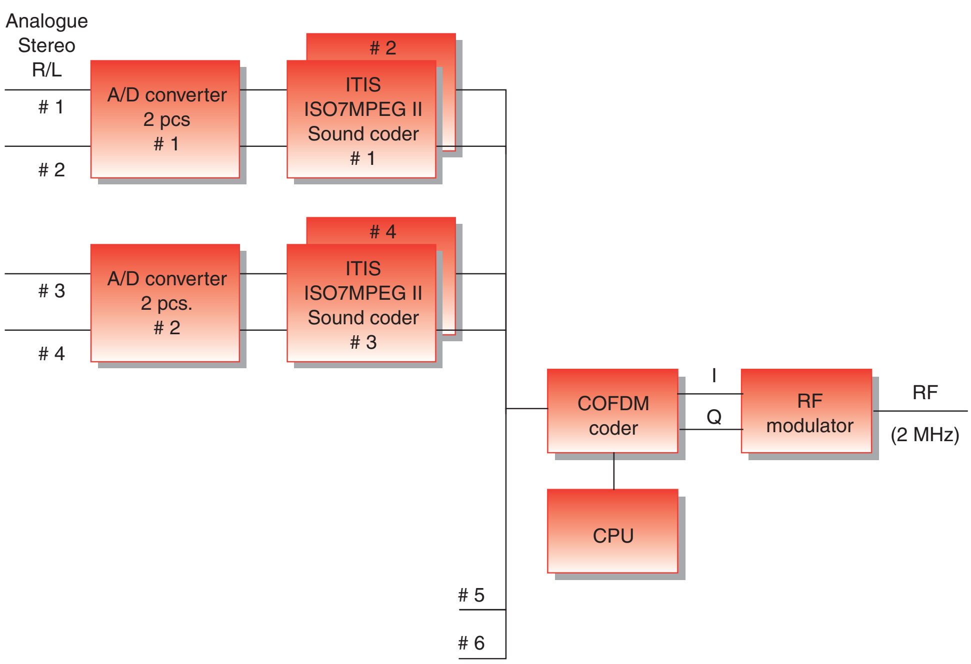


Figure 2 DAB encoder

Frequency aspects

The frequency planning situation in Norway is similar to most of the other countries in Europe. In the first years of the planning study it was assumed that we would use the band 100 – 104 MHz, but after the introduction of national radio stations P3 and P4 this has become impossible. It is possible that DAB in Norway will use frequencies in the high end of Band III, TV channel 12 and up to 240 MHz. The test network has a licence from the Norwegian Telecommunication Authority to use frequencies between 235.25 and 240.00 MHz. The test transmissions will use 236.75 ± 0.75 MHz.

The DAB test network

It was decided to install DAB transmitters on two main transmitter sites; Tryvann, Oslo and Jonsknuten, Kongsberg. There is also equipment for a third smaller transmitter, which can be installed when the first preliminary tests show where it would be most practical. The most likely position is Hvittingen near Sande, south of Drammen.

Programme feed

The MPEG layer II encoder and COFDM modulator are situated at NRK in Oslo. These comprise an ITIS encoder, and a COFDM and RF modulator from Telefunken. The encoder was originally ordered with two stereo audio channels. It will be expanded in the near future to four stereo channels. The output of the modulator is downconverted to 2 MHz and transmitted to the DAB transmitter on an analogue video connection (cable).

The traditional way of feeding main transmitters is by microwave link. The fill-in transmitters in this case are fed by an “off-air” signal which is then frequency converted. In a DAB single frequency network (SFN) this is impossible. The repeater has to transmit the same signal on the same frequency as the input signal. Calculations and tests show that the isolation between receiving and transmitting antennas is less than the required gain. The SFN also requires identical data signals on adjacent transmitters. The transmitting frequencies must also be very accurate, less than 10 Hz difference is tolerated.

The remote stations will be fed with UHF links. The signal sent to Kongsberg/Jonsknuten is transmitted on channel 40 using a modified TV transmitter. Synchronization is achieved by using the carrier as a pilot frequency and the COFDM signal as the upper sideband. The output stage in Oslo/ Tryvann is a standard 20 W UHF power amplifier.

DAB transmitter (VHF)

The main transmitter site is Tryvasshøgda radio which is an unmanned station. The main tower at Tryvann carrying TV and FM antennas could not be used due to lack of space. The DAB transmitter is a Hirschmann VHF/TV transmitter, with a TH298 tube producing 1 kW. The signal transmitted from Oslo/Tryvann must be delayed by 245 µs to compensate for the distance of the UHF link to Jonsknuten/Kongsberg, about 74 km. The input up-converter is the same type as used for the German satellite system, and converts the signal via a standard IF of approx. 38 MHz and a standard TV up-converter. The 1 kW DAB transmitter set comprises two 19" racks. One rack houses the input module, converter and frequency reference system, and the other contains the 1 kW tube amplifier. The DAB signal requires a linear amplifier. The ratio between peak power and average power is more than 10 dB. It is assumed that intermodulation products just outside the channel below -35 dB are acceptable, but there is no internationally agreed limit. When linearity correction is used in the power amplifier, it is possible to attain 1 kW DAB power with intermodulation products approx. 35 dB below the output level.

The DAB transmitter at Kongsberg/ Jonsknuten is similar to the Oslo/Tryvann transmitter, except that this station also has a UHF down-converter. The same unit also obtains synchronization from the pilot carrier, thus ensuring that both VHF transmitters are running at the same frequency.

Transmitting antenna system

When installing a new network, one important consideration is the antenna type used. It is normally assumed that vertical polarization gives best reception for mobile applications with a whip antenna. The most difficult areas to obtain coverage in practice are those places in the shadow of the direct beam from the transmitters. The signals reaching such areas are reflected or refracted. In such cases the polarization may also be changed. The received signal level in those places may not be much different whether a vertical or horizontal transmitting antenna is used. If horizontal polarization could be employed, it would be possible to use the high gain antennas for TV VHF transmission. The coverage test with both vertical and horizontal polarization is therefore of treat interest.

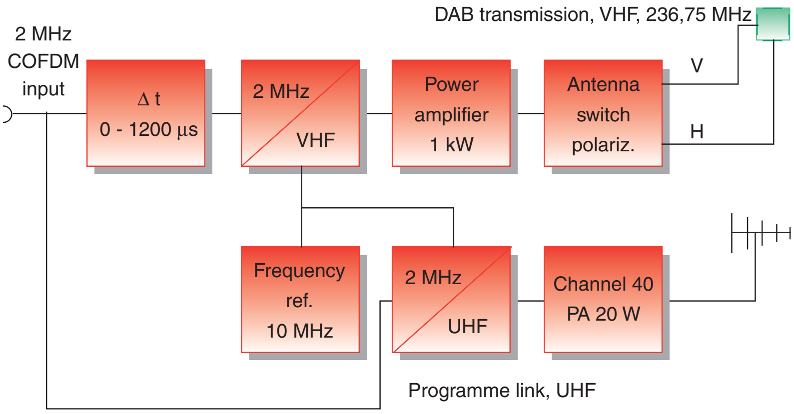


Figure 3 Transmitter equipment, Tryvann

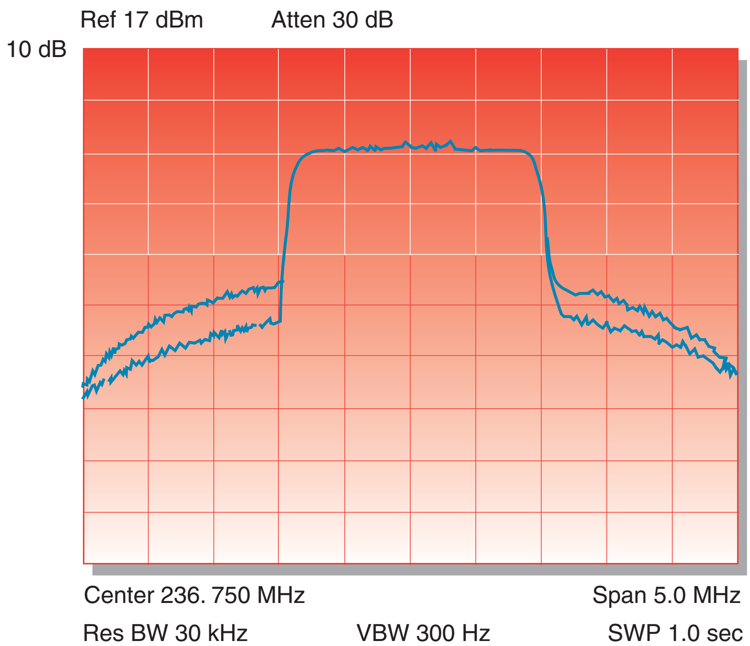


Figure 4 Intermodulation in a 1 kW tube amplifier.
Upper curve – without linearizer

Oslo transmitter, Tryvatn

Three antenna panels are installed, type Kathrein, with both horizontal and vertical polarization.

Height above ground:	110 m, and 640 m above sea level
Antenna gain:	approx. 3 dB
Feed cable:	Kabelmetal, HF 1 5/8".

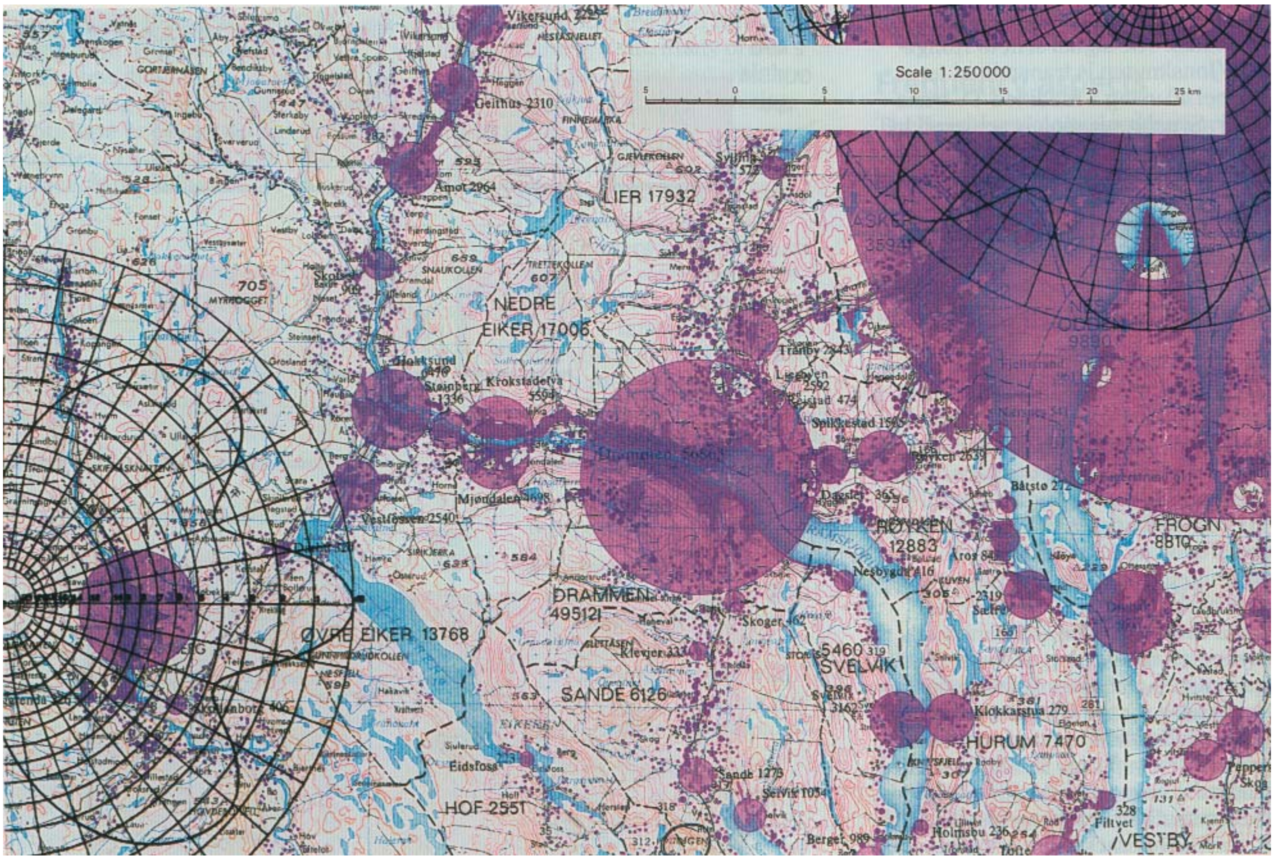


Figure 5 Antenna diagrams and population in coverage area

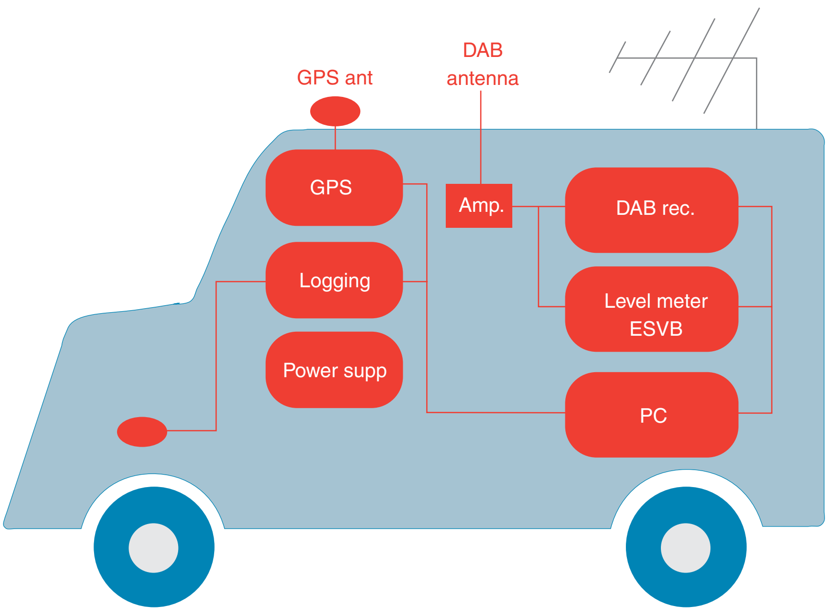


Figure 6 Measuring van

The vertical and horizontal dipoles are fed by two different feeder cables. The switch between the two antenna systems can therefore be remotely controlled.

With a transmitter power of 1 kW, the ERP will be approx. 1.5 kW between 70° and 300° N.

Kongsberg transmitter, Jonsknuten

Two antenna panels of the same type as Oslo are situated at the top of the FM/TV transmitter station tower.

Heigh above ground: 91 m, and 898 m above sea level.

With a transmitter power of 1 kW, the ERP will be approx. 2.5 kW in a sector 20° – 170° N.

Receiver and measuring equipment

Eureka 147 3rd generation receivers will be used by NRK, Norwegian Telecom Research and Broadcasting. The receiver

Table 1 Necessary DAB field strength calculation

Required C/N with interference allowance	10.0	dB
Receiver input thermal noise (1.5 MHz BW) 75 ohm	-3.5	dBµV
Receiver noise factor (conservative estimate)	10.0	dB
Antenna coupling loss, whip antenna at 230 MHz, approx.	10.0 +	dB
Required field strength at antenna position	26.5	dBµV/m
Correction to refer 2 m field strength to 10 m above ground level	12.0	dB
Location correction factor	13.0	dB
Allowance for man-made noise	4.0 +	dB
Calculated median necessary field strength, 10 m above ground level	55.5	dBµV/m

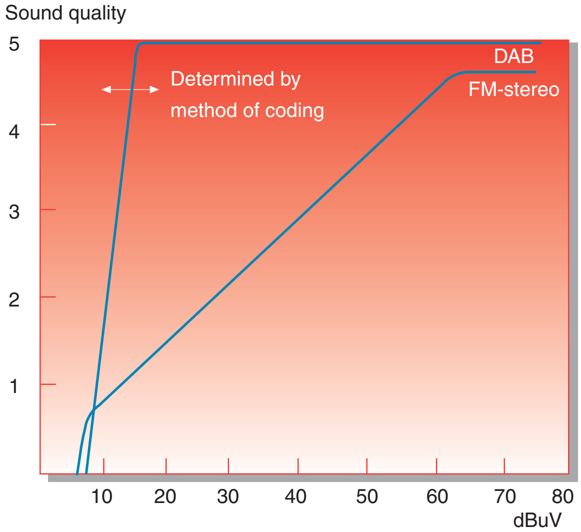


Figure 7 Audio quality versus input signal for DAB and FM-stereo

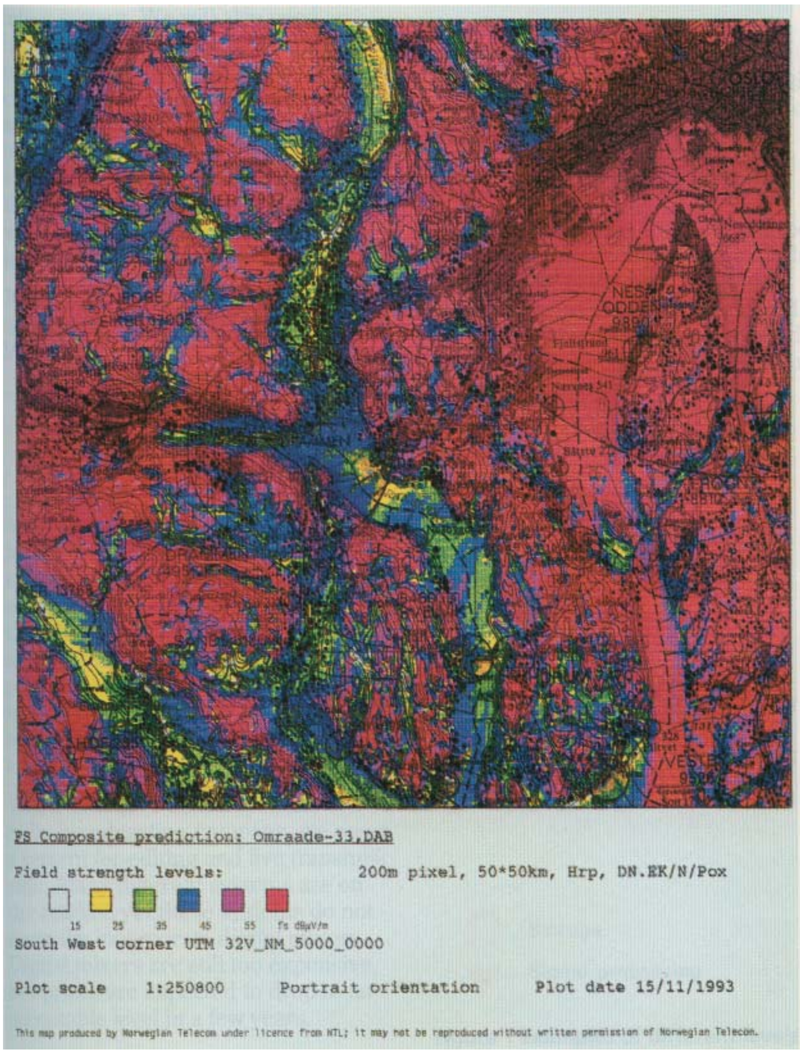


Figure 8 Coverage prediction, a part of DAB network

weighs about 10 kg and has a power consumption of 50 W from 12 V d.c. It can receive on bands 1, 3, 4, and 5, and has a system IF bandwidth of 1.5 MHz. The receiver can decode one of six stereo programmes per channel. The receivers have been tested together with the transmitters and the whole chain is now functioning. We were warned about spurious RF signals from the receiver. Several frequencies are radiated at relatively high levels, but luckily enough, none of them is in the frequency band that will be used for transmission in Norway. We will modify the receivers to reduce radiation.

The measuring van will contain one DAB receiver, one field strength meter (Rohde & Schwarz, ESVB), a band III whip antenna and a calibrated measuring antenna. The positioning system will consist of a GPS receiver together with a module which has a compass and an interpreter which receives impulses from the wheels. The system will be controlled by a PC which will also log all signal measurements. The output file will contain the following information: Signal strength (from the ESVB), signal quality (from the DAB receiver) and position (from GPS). The data will be used to derive a measured coverage map and this will be compared with the calculated coverage map.

Field strength considerations

Figure 7 illustrates a very important point. The signal level necessary for *high* quality sound is much lower for DAB than for FM-stereo. However, the level for FM that gives mono signals with some noise, is approximately the same as for viable DAB reception. For mobile reception (car radio) this level is the “low usable level” for FM as the signal level is only so low a relatively small percentage of the time. The same signal level for DAB gives full quality, but has a very steep cut-off curve. The “low usable level” for DAB has to be much higher to allow for field strength variations. A calculation for field strength must be carried out for 99 % of locations instead of 50 % which would be used for FM. The DAB location correction factor for 50–90 % of locations has been measured[1] to be approx. 13 dB. This is considerably less than the FM location factor for 50–90 % (19 dB).

Table 1 shows a first and probably conservative estimate. Network gain of a Single Frequency Network, with more than one transmitter would give a reduction of 0 – 8 dB. Both this figure and necessary field strength will be verified in the mobile and stationary field tests.

The calculated field strength map for the Drammen–Oslo test area (Figure 8) is composite calculation. Only the highest level of field strength from either of the two transmitters is shown. Network gain is not calculated. This map shows a very interesting area around Drammen and Lier where the calculated field strength values vary around the calculated low level limit.

The height profile between Tryvann and Jonsknuten is shown in Figure 9. The Lier valley is situated approximately 1/3 in from the left hand side. The profile shows clearly that there is no direct path to any of the two transmitters in this area.

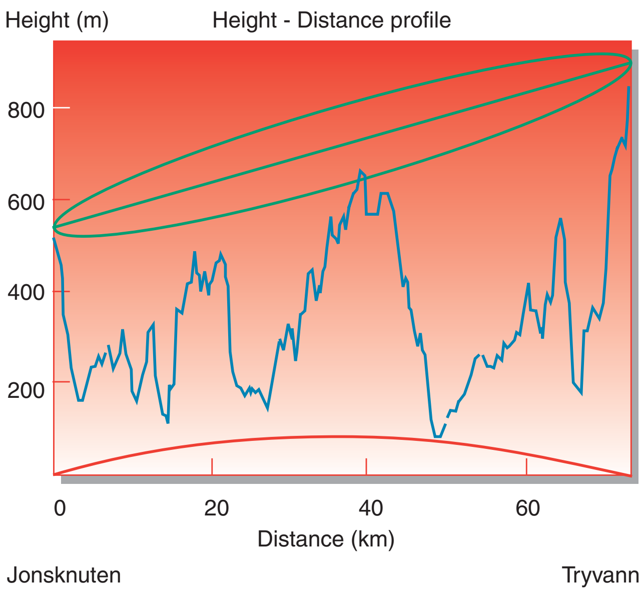


Figure 9 Height profile, Tryvann - Jonsknuten

---

[1] *BBC RD 1993/11 Digital Audio Broadcasting*

Radio production – evolution towards an all digital environment

BY ARE NUNDAL

Introduction

As the evolution of audio is moving into the digital domain it is becoming clearer that the professional audio industry is no longer capable of developing the components needed in radio production. The cost involved in developing the components required in modern professional audio equipment, is so large that few manufacturers would be able to get the equipment to the market with reasonable returns on the investment. The volume of units that can be expected to be sold is far too small.

Luckily, the computer industry supplies us with most of the necessary components needed to design and produce professional digital audio equipment. The components we need are: storage media, signal processing devices (including converters), transfer networks, and a platform on which these components can be assembled into a system. We will also need an interface to enter control information from the user. One can, quite rightly, claim that this is not much different from other complex computer controlled systems, e.g. process control for chemical plants. The main point is that the components are already there, with one possible exception, the user interface.

When developing a system for digital audio we need to address the following topics:

- System integration
- User interface
- Standards for digital audio.

Most importantly; the equipment must have a price/performance ratio that can compete with existing analogue equipment. Most broadcasters will claim that increased sound quality alone cannot justify any increase in price compared to analogue equipment. The advantage of introducing new systems has to give a more efficient production of radio programmes.

When talking of digital systems in a radio broadcasting company, we mean a system for editing and live transmission. Many different systems are on the market, yet these systems do not normally incorporate digital mixers. Digital mixers are still too expensive, but prices are expected to drop to an acceptable level in a few years.

In a medium or large radio broadcast station, there are some basic constraints on the digital system for audio.

- The system must allow for multiple users at any time.
- The users will be working at separate workstations which must be able to access a central database of audio files.
- All of the users must to be able to put their edited audio items into a common playlist or central archive.
- The users must be able to record, edit and transmit the different audio files from anywhere in the system via the data network.
- The users must have access to other services on their workstations, i.e. word processing, news agencies, databases and scheduling systems, to mention a few.

In the following discussion we will examine some of the points arising when designing and purchasing digital systems for professional radio production.

Sharing the resources?

The first question is where to put the storage resources in our audio production system? Is the best solution to have one large shared storage for all the audio items? This storage would have to accommodate all the files containing the audio items in the system. The audio has to be distributed back to the workstation during editing and playing via the data network or along analogue lines. Should we alternatively place the storage on each workstation? Each user could put his or her audio files on their personal workstation, while the common audio files are placed on a central workstation for live transmission. Network database techniques would have to be used to keep track of which audio files and which versions are kept on the different workstations.

Having decided where to put the storage resources, we have to consider the placement of the signal processing resources. There are again several possibilities, but in this type of system we can assume that we would not have distributed storage while having centralized signal processing resources. Therefore, if we have distributed storage, we would also have distributed signal processing resources. If, on the other hand, we have centralized storage, we could choose whether to have centralized or distributed signal processing. The three possibilities are shown in Figure 1.

In the first case we would have to take some kind of control information from the user terminal back to the central system, and we would also need to send the audio signal on analogue lines from the central system to the user terminal. The initial cost of this kind of system is high,

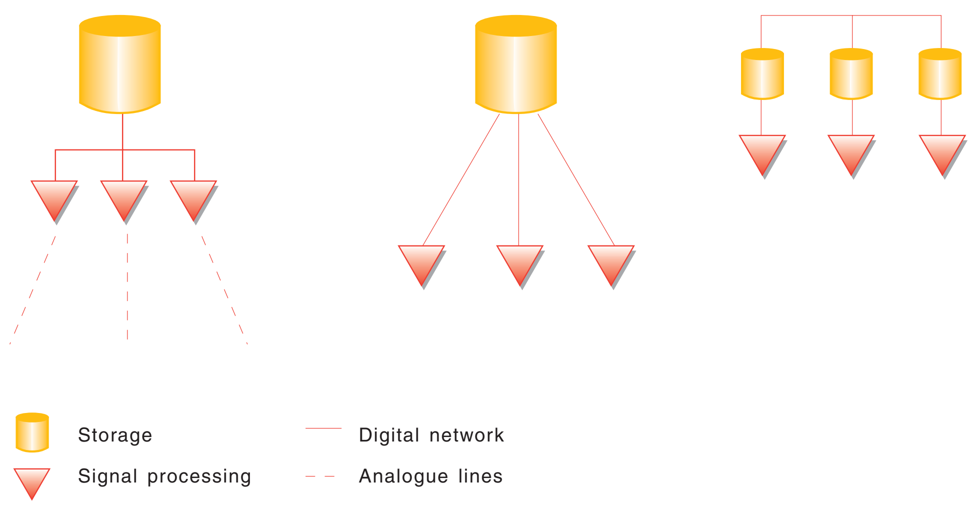


Figure 1 Examples of different levels of distribution of storage and processing resources

but such systems are able to handle a large number of simultaneous users, even with linear PCM coded audio. Additional user terminals could also be added at lower costs compared to the other two system alternatives.

The second alternative is to have a central storage for audio, while having the signal processing and conversion hardware at the user terminals. In this case we would have to transfer control information between the user terminal and the central system, but we would also have to transfer audio, as digital data between the user terminal and the central system. For these two purposes we use a data network. Most commercial data networks could be used. The number of simultaneous users is limited by the capacity of the network and the data rate at which the audio is coded. In order to be cost effective with current technology, these systems would have to use low bitrate codings. We will discuss some aspects of low bitrate coding later in this article. In real life, the user terminals would be ordinary PCs with plug-in cards, and the network would be of Ethernet or token ring type. In order to increase the number of simultaneous users beyond approximately 20, we would have to use more network segments.

The last alternative would be to distribute both storage and signal processing. The user terminals would be ordinary PCs with plug-in cards and extra storage, connected together via data networks. In order to make such a system work as a total sound production system, we would have to use distributed database mechanisms. This solution has up to now only been realized for high cost/performance systems.

Compressed or linear audio?

One of the questions that always arises when we are talking about digital audio is, how can we reduce the amount of data that is needed to store and transfer the audio. Several standards have already been defined for reducing the data rate of high quality digital audio. Most of these standards use psycho-acoustic principles of human hearing to remove data without altering the perceived sound. Tests have shown that even with very significant reduction of the datarate, even trained listeners are not able to tell the difference between the original and the bitrate reduced sound. The coding is perceived as being transparent. Table 1 gives an example of bitrates for coded audio giving acceptable quality.

The two standards of bitrate reduction quoted in this table are accepted as giving transparent quality for most types of audio material.

There are some drawbacks of bitrate reduction. Audible side effects can be produced for some kinds of complex music but this signal degradation is not usually serious. Signal degradation is more severe when bitrate reduction coding is used in series (tandem coding). This degradation can be made audible by processing which alters the spectral balance or the dynamic properties of the material between the codings. Encoding/decoding without intermediate processing 5–8 times, is accepted to be the upper limit whilst maintaining transparent sound. A delay is introduced in the encoding/ decoding process. This is due to the size of the block containing coded samples. The smallest unit in the bit stream is no longer the sample, but the 'block'. This also limits the accuracy of the edits done with the encoded sound. It is very difficult to perform traditional signal processing on the sound without decoding it, e.g. mixing and equalizing.

All of these factors in some way restrict our use of bitrate reduced audio. It is very easy to show that in a radio broadcasting chain, from microphone to receiver, the assumed maximum number of codings is often exceeded. We will therefore have to restrict our use of these codings. To decide which parts of the broadcasting chain to reserve for bitrate reduction codings, we will divide the chain into four categories or parts.

- Contribution
- Production
- Distribution
- Transmission.

When examining these parts, and comparing the cost involved, it is evident that the most expensive part is in transmission. Radio frequency bandwidth is a very restricted resource. It is obvious that audio bitrate reduction coding must be used in order to keep the usage of this resource to a minimum. This is reflected in the DAB (Digital Audio Broadcasting) standard.

When comparing costs for disk space for storing audio with transmission, the majority of the costs are involved in the transmission. We should also therefore give priority to the contribution and distribution part of the broadcasting chain when considering bitrate reduction coding techniques. This does not necessarily mean that we should avoid bitrate reduction in production systems, but when we approach the limit for tandem codings, production is definitely the part of the chain where we should turn to linear audio. We have to take this into account when designing and specifying production systems.

User interface

Before looking into the details of the user interface we have to examine two factors concerning the users in larger radio broadcasters:

- One aim is to involve more professions in editing and live production than the traditional engineers.
- The potential user group spans a wide range, considering age, knowledge and ability to adopt to new technology.

It cannot be emphasized enough: If all the users mentioned above are to use the system extensively, it must be *simple.* Simple in this context is very difficult to define, but there are a few rules of thumb:

Table 1 Data rates on the currently most common data compression schemes used on PC based editing systems

Format	Data rate	Bytes/min	1 Gbyte =	Reduction
Coded ISO/MPEG layer 2	2 x 128 kbit/s	1,92 M	8h20	6x
Coded ISO/MPEG layer 2	2 x 192 kbit/s	2,88 M	5h43	4x
Linear	2 x 768 kbit/s	11,52M	1h27	0x

- The process must resemble a familiar process, e.g. the splicing of tape. Jog-wheels resemble turning of tape wheels, faders have a natural association with levels, etc.
- Specialized user panel with dedicated push buttons, fader and knobs where each control has its own unique function, would be preferable to general user interfaces like the keyboard and mouse.
- The last item leads us to another major concern; ergonomy. The user interface has to be comfortable to use. It must not be tiring on any part of the body.

If one follows these guidelines an intuitive user interface may be realized. The function of each control will be obvious to the users, and simple operations are thus easily performed. If the interface is intuitive, it will be easy to introduce the new technology to established users. The equipment can be used effectively from the beginning, and the need for training can be reduced to some extent. In other words, a good user interface is vital when introducing new technology.

Production formats

Production studios are often equipped with a swarm of different production formats. The equipment for the old analogue formats like 1/4" tape, vinyl records and compact cassettes have not been rendered unnecessary, while the digital formats CD and DAT are well established in the control room together with computer hard disc based systems, analogue Carts and the new digital floppy disc based carts. It is by no means economical to have all these production formats in use. We have to reduce number of formats in the future.

The natural question to ask is then, which formats to keep and which of them will exist in the future. Guessing is very difficult and many of those who have tried to predict the future have failed. Nevertheless, Table 2 gives us a starting point by summarizing some important qualities of the different professional formats.

The analogue formats will eventually disappear. DAT and CD are commercial formats. CD has proven to be a very good way to distribute pre-recorded material. DAT will eventually be replaced by some of the other digital formats.

Surviving formats for production must give very fast access to the material. Formats with random access are therefore the only choice. It is possible that all three random access formats will prevail. Hard discs because they are cheap and an easy way to store relatively large quantities of audio. MO discs because they are easy removable, and can withstand relatively rough handling. Solid state RAM because of its extreme robustness which is well suited for portable recorders. Cost and long term stability are most important for archiving purposes. It is therefore possible that the QIC cassette can survive as a format for long time mass storage of audio.

Add to these formats possible new formats that technological development will bring. The storage capacity of optical tape is only one example of the advantage of these new formats. Other formats will surely follow in the future.

Table 2 A short form comparison between the most common production formats used in professional broadcasting companies

Format	Access	Handling	Price
1/4 tape	linear/ cut	rough	
Vinyl LP	linear	fragile	
CD	random	rough	high
DAT	linear, no splice	fragile	medium
QIC	linear, no splice	rough	low
MO disc	random	medium	medium
Hard disc	random	fragile	low
RAM	random	rough	high

Standards

Organizations like EBU, AES and ISO have given us some very useful standards. This is a necessary condition for the broad acceptance of the technology. Here we will only mention a few of the most used and accepted standards.

AES/EBU interface. This standard specifies the point-to-point interconnection of linear coded (PCM) audio between two pieces of equipment. The standard is only intended for connections over short distances. The AES/EBU interface is implemented on virtually all professional digital equipment. It is the de facto industry standard. Recently, there have been attempts to extend the standard to incorporate five general bit streams, each of which can be used for bitrate reduced audio.

ISO/MPEG has defined a standard for coding of bitrate reduced audio. The standard is defined in three different layers of complexity. Within each layer there are several bitrates available. This format is predominant in Europe. See the section treating this topic separately.

MUSIFILE, is a file format standard for digital audio, both linear and bitrate reduced. With this format implemented it is possible to use different products within a system for radio production.

These are a few of the standards that exist for digital audio, but definitely the most important ones, excluding CD and DAT. We also have to rely on several standards in the computer industry. Standardization is a key point when products from the computer industry form the basic components in radio production systems. As radio broadcasters, we should contribute to and encourage the standardization work that is going on in the different organizations.

An example

To end this article we shall present a system that is installed in NRK's district office in Oslo. This office is a medium sized radio station with a focus on local news and current affairs. A digital production system was required for their editing and live transmissions. Tandberg data was chosen as the contractor, and simplicity of the user interface was emphasized. The system comprised eight workstations in a network with a central server, see Figure 2. The workstations are regular PCs with an add-on card containing signal processing and conversion hardware. A dedicated hardware control

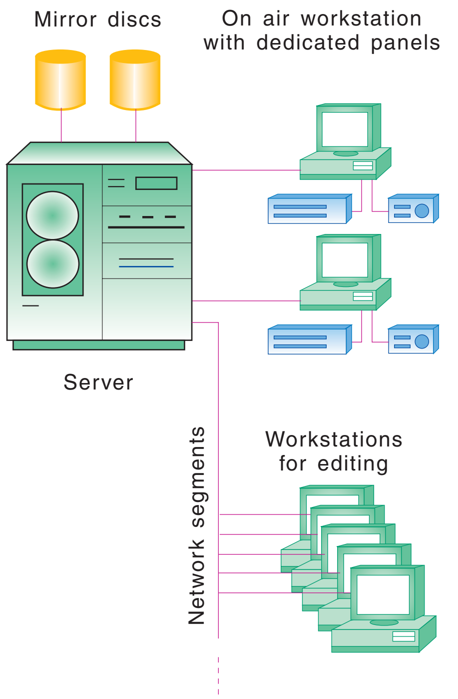


Figure 2 Schematic view of the system for digital production and on air broadcasting, installed at NRK's district office in Oslo

panel is used to simplify the use of the system, see Figure 3. Bitrate reduced audio is used according to ISO/MPEG layer 2 with a bitrate of 192 kb/s per channel. The central storage has a capacity of 20 hours of audio.

The system is integrated with the standard news service. Although much emphasis was put on simplicity of use, the system was initially not simple enough, and was not accepted. The contractor rewrote parts of the program and a few months later the system was accepted by the users, and it is now in use. Figure 4 is an example of the screen layout when editing.

Conclusion

When equipment for radio production is assembled from components from the computer industry, it is very important that we involve the users in the design phase. It is easy to forget that we are not making computers, but equipment for producing and playing audio on the radio. Perhaps future generations will prefer modern computers to dedicated equipment, but in the meantime we have to remember the purpose for which the new equipment is to serve: to help all the users make better programmes, faster.

Figure 3 Layout of the dedicated user panel for editing on the Tandberg digital audio production system

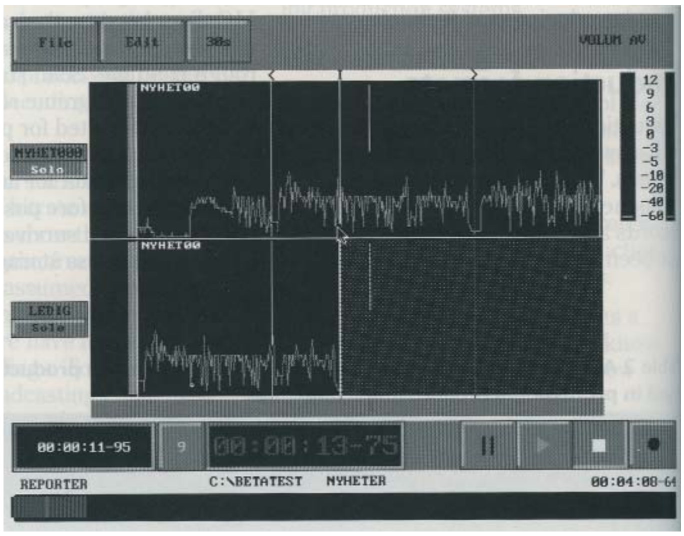


Figure 4 Screen layout of the editing programme on the Tandberg digital audio production system

Television production – evolution towards all digital systems

BY PER BØHLER

Introduction

It is fair to claim that the change from analogue to digital technology in television production is the greatest technological change since television was introduced. When the first black and white television was replaced by colour television, it was an evolutionary and compatible step on the ladder of analogue television.

The introduction of a digital technology represents revolution in its own right, there is no link to the analogue world except via A/D- and D/A converters. As the video signal is replaced by binary digits, equipment and systems will also have to be replaced. Where analogue and digital systems have to work side by side in an interim period, conversions between the two worlds have to take place.

Today it is probable that NRK will replace the analogue infrastructure with digital systems towards the end of this century. This is provided our economy will allow the necessary investments.

Modern analogue TV production

When colour television was introduced in Norway in 1972 it represented one of the last steps of analogue television systems. The PAL system, in concept a transmission system eliminating the drawbacks of NTSC system, was also well suited for use in studios as a production system.

The PAL composite colour coded signal can easily be handled by the production equipment provided the necessary precautions are taken regarding synchronization and timing. As the videotape recorder (VTR) improved and electronic editing systems matured, one inherent deficiency in the PAL signal became a problem. Originally, the absolute phase between the colour sub-carrier and the sync signal was not defined in the PAL system. This created a timing error when editing videotapes containing signals from different sync sources. To overcome this the PAL signal specifications were modified and a fixed relationship between sync and the colour sub-carrier was introduced.

Even with this modification the PAL signal is not easy to handle for the VTR in the editing phase. To explain this briefly, let us look at the equation describing the relationship between system line frequency and colour sub-carrier frequency.

$$f_{SC} = (567 / 2 + 1 / 4)\, f_L + 25 \text{ Hz}$$
$$= 283.75 \times 15625 + 25 \text{ Hz}$$

$$f_{SC} = 4.43361875 \text{ MHz}$$

f_{SC} = subcarrier frequency

f_L = line frequency

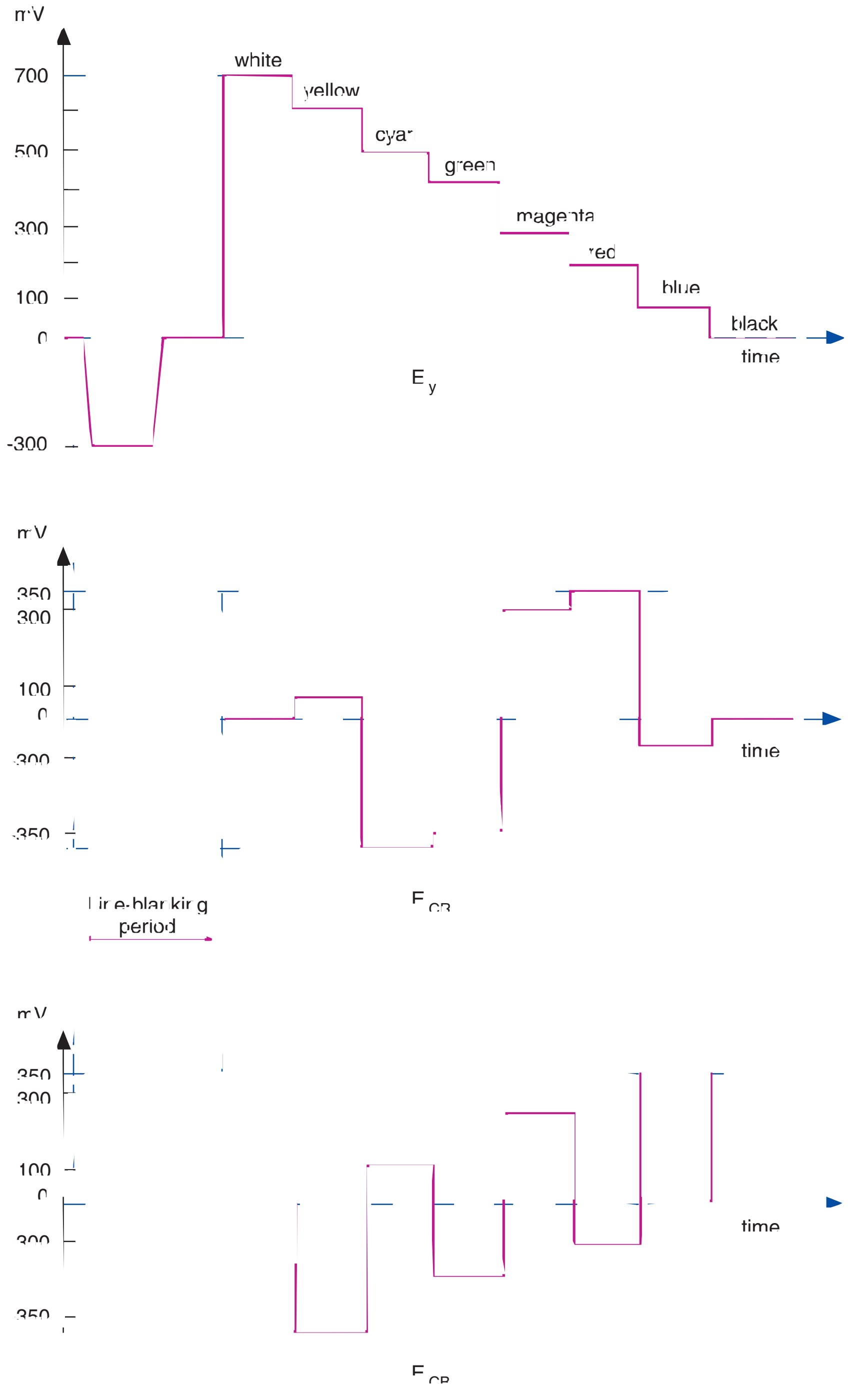


Figure 1 EBU N10 signals

One can see from this relationship that the sub-carrier frequency contains an integer number of cycles for every eight fields. This means that in an editing process there is 160 ms between edit points; a considerable time interval when there is much action in the scene. In addition, the VTR will have to identify one field out of eight when synchronizing to the studio reference, a process taking a few seconds every time the machine starts.

Table 1 Encoding parameters for Rec. 601, 4 : 2 : 2 member of the family

Parameters	525-line 60 fields/s systems	625-line 50 fields/s systems
1 Coded signals: Y, C_R, C_B	From gamma pre-corrected signals	
2 Number of samples per total line: - luminance *(Y)*	858	864
- each colour difference signal (C_R, C_B)	429	432
3 Sampling structure	Orthogonal, line, field and frame repetitive. C_R and C_B samples co-sited with odd (1st, 3rd, 5th, etc) Y samples in each line	
4 Sampling frequency[1]: - luminance signal	13.5 MHz	13.5 MHz
- each colour difference signal	6.75 MHz	6.75 MHz
5 Form of coding	Uniformly quantized PCM, 8 (optionally 10) bits per sample, for the luminance signal and each colour difference signal	
6 Number of samples per digital active line: - luminance signal	720	
- each colour difference signal	360	
7 Analogue-to-digital horizontal timing relationship: - from end of digital active line to OH	16 luminance clock periods	12 luminance clock periods
8 Correspondence between video signal levels and quantization levels:	(Values are decimal)	
- scale	0 to 255	
- luminance signal	220 quantization levels with black level corresponding to level 16 and the peak white level corresponding to level 235. The signal level may occasionally excurse beyond level 235.	
- each colour difference signal	225 quantization levels in the centre part of the quantization scale with zero signal corresponding to level 128.	
9 Code-word usage	Code words corresponding to quantization levels 0 and 255 are used exclusively for synchronization. Levels 1 to 254 are available for video.	

[1] The tolerance for the sampling frequencies should coincide with the tolerance for the line frequency of the relevant colour television standard.

The PAL-system therefore makes high quality editing difficult and time consuming and in addition puts limitations on the achievable picture quality due to:

1 intermodulation distortion in the recording process caused by non-linearities,

2 folding of sidebands caused by the choice of modulator carrier frequency and deviation as defined in the recording standard.

To overcome these problems recording systems without use of a high level sub-carrier and with separate recording channels for luminance and chrominance information was developed. The most successful of these system have been the Betacam analogue component recorder. These recorders have been sold world wide in more than 180 000 units. In addition, other system components have also been developed and analogue component production systems are today giving superior picture quality compared to PAL studio systems although the signal is encoded to PAL for transmission.

The signal format for component systems is internationally standardized as in the EBU N10 standard for the European 625 line system. A similar standard also exists for the 525 line system. Signal wave forms and amplitudes are shown in Figure 1.

The early years of digits in television

In the early seventies the VTR had a very limited correction window in the Time Base Corrector, only 1 µs approximately. This was caused by the fact that the video signal had to be stored in analogue delay lines with variable delay. VTRs had been built with ultrasonic glass delay lines in the TBC giving a correction window of one TV line. However, these machines were very expensive and complex; in addition the video buffer suffered from the general drawbacks associated with ultrasonic glass delays.

To overcome these problems the last generation of 2" Quadruplex VTRs introduced a digital TBC with a correction window of more than half a TV line. The digital memory could store 512 bytes! At that time it was quite an achievement. But by today's standard the digital process was quite simple; the A/D and write processes were clocked by three times subcarrier frequency from the playback signal. The read process and D/A conversion were locked to studio reference, thereby providing a stabilized output signal.

Around the mid-1970s the first generation of 1" helical scan VTRs with broadcast quality came on the market; these machines improved rapidly in user flexibility and performance as the use of digital technology increased. With the correction window of the TBC increased to many TV-lines, the start-up time could be shortened considerably. This was of great importance in order to reduce the time needed for post production editing. Towards the end of the 1970s increasingly complex post-production processes started to appear, and this demanded access, not to the composite signal, but to its components. It also demanded a sampling structure that was repeated exactly from one picture to the next, which is not the case with sampling clocks locked to the colour sub-carrier. It was also fortunate that work carried out had established that it is possible to interface a component coding system with a composite PAL environment using a digital PAL decoder with the sampling clock not locked to PAL sub-carrier.

At the same time better designs of A/D and D/A converters were developed which, when applied to composite signals, were able to use sampling clocks not locked to sub-carrier without producing unwanted patterning in highly saturated areas.

Towards an international coding standard

On this background it was clear that if the international discussions should focus on component coding, the choice of sampling frequencies would not have to be restricted by the need for a "composite/component bridge" close to a multiple of the sub-carrier frequency.

A period with intense international discussions followed to search for acceptable frequencies. The concentration was now on component signals with an emphasis to seek for the greatest possible commonality between 525- and 625-line applications. The result was the well known CCIR (now ITU/R) Recommendation 601 which defines the coding parameters for both 525- and 625-line systems [1]. This recommendation in fact describes a whole family of encoding standards where the 4:2:2 member is most common. For both systems the same sampling frequencies are used (13.5 MHz and 6.75 MHz for the Y and U/V signals respectively) and the number of samples per active digital line are also the same (720 samples and 360 samples for Y and U/V respectively). The number of samples per total line are different, but these are integers (858 for 525-line and 864 for 625-line) since the sampling frequencies are multiples of 2.25 MHz.

ITU/R Recommendation 601

Rec. 601 is the coding standard describing all parameters and details related to the digital encoding of component video signals. The main characteristics are given in Table 1.

In the table the colour difference signals C_B and C_R are renormalized *(R-Y)* and *(B-Y)* signals. The explanation is as follows:

In analogue (and digital) form the luminance signal *(Y)* has a relative amplitude of 1. The colour difference signals *(R-Y)* *(B-Y)* have relative amplitudes of ±0.7 and ±0.89, respectively. In order to give the digital colour difference signals the same relative amplitude as the luminance signal, the following re-normalization coefficients will be:

$$K_R = 0.5/0.7 = 0.71$$

$$K_B = 0.5/0.89 = 0.56$$

The expressions for C_B and C_R then become:

$$C_R = 0.71(R\text{-}Y) = 0.50R - 0.42G - 0.08B$$

$$C_B = 0.56\,(B\text{-}Y) = -0.17R - 0.33G + 0.50B$$

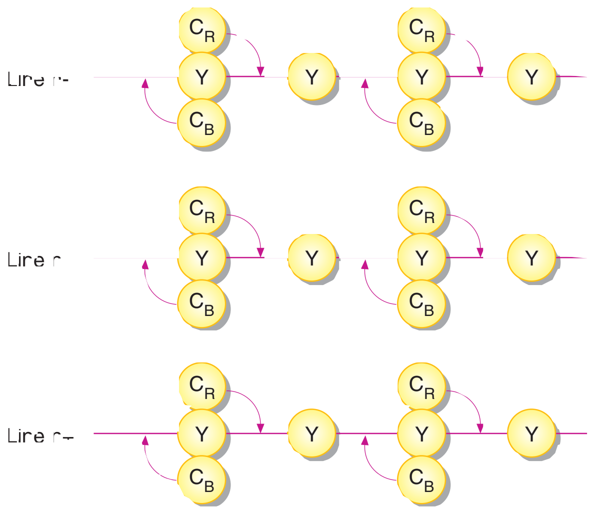


Figure 2 Orthogonal sampling structure and co-sited samples

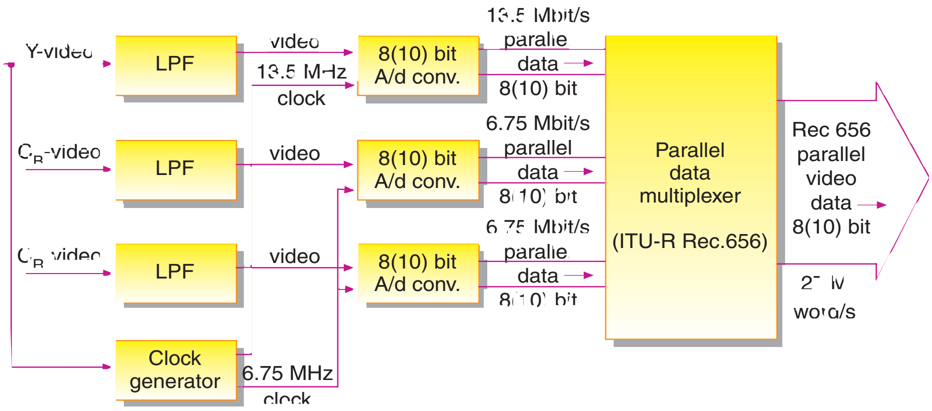


Figure 3 A/D converter according to ITU-R Rec. 601

Specification for a luminance or RGB signal filter used when sampling at 13.5 MHz

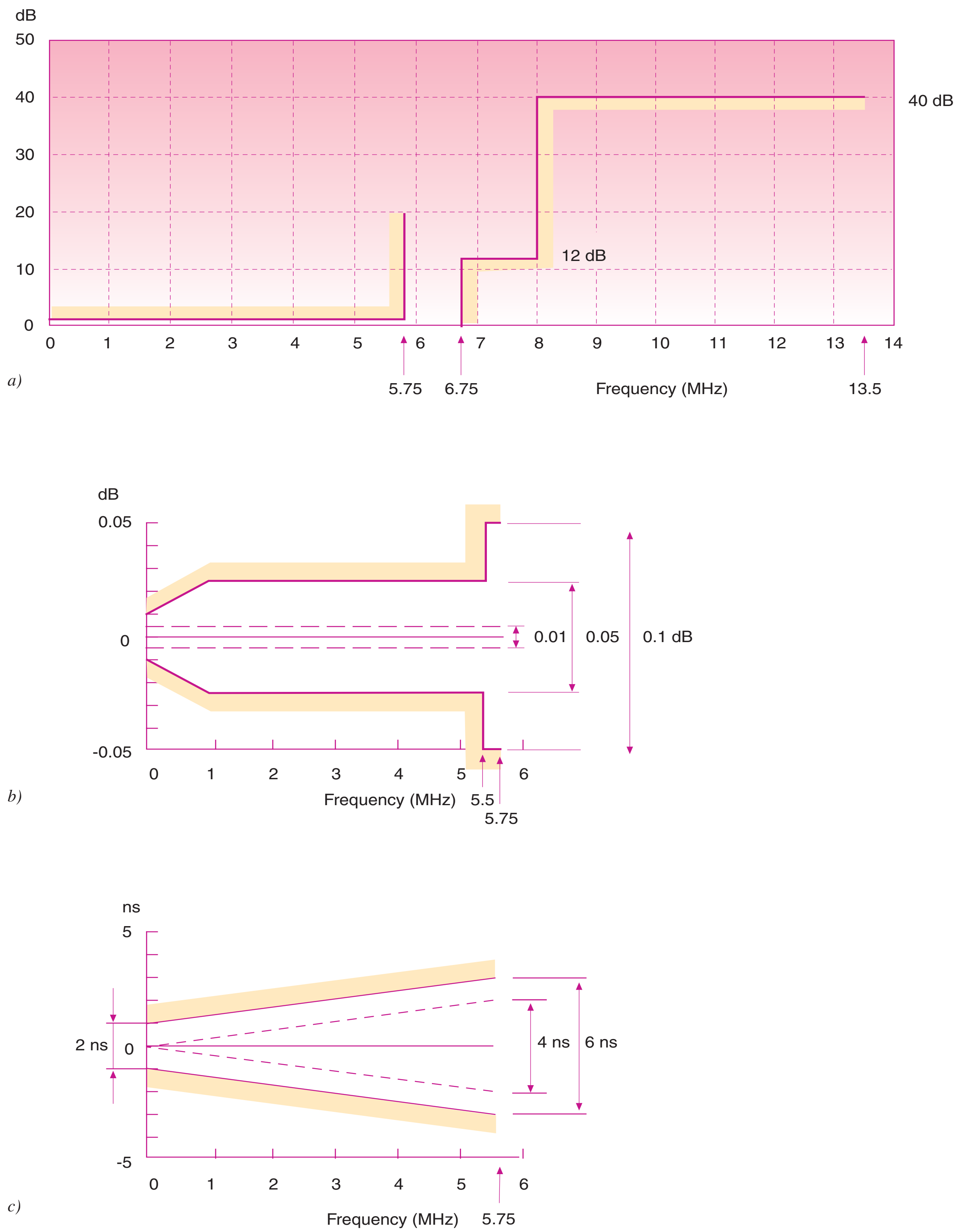
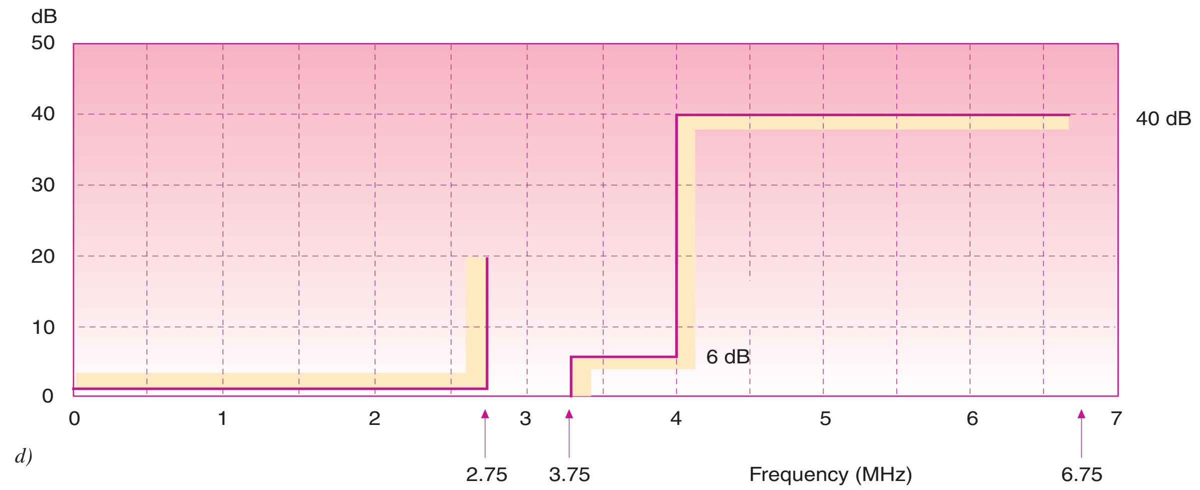
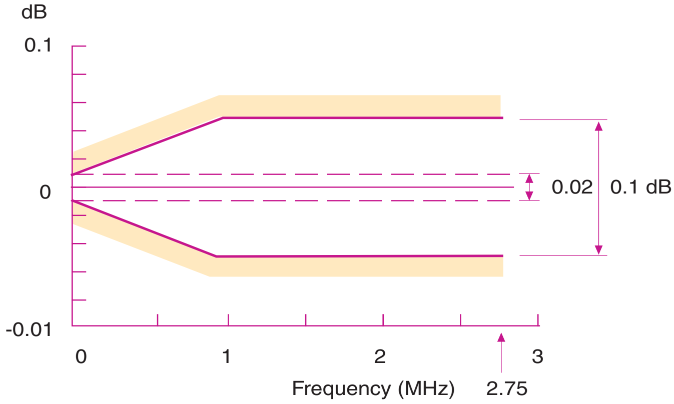


Figure 4 Anti-aliasing and reconstruction filters according to ITU-R Rec. 601
a) Template for insertion loss/frequency characteristic
b) Passband ripple tolerance
c) Passband group-delay tolerance
Note: The lowest indicated values in b) and c) are for 1 kHz (instead of 0 MHz)

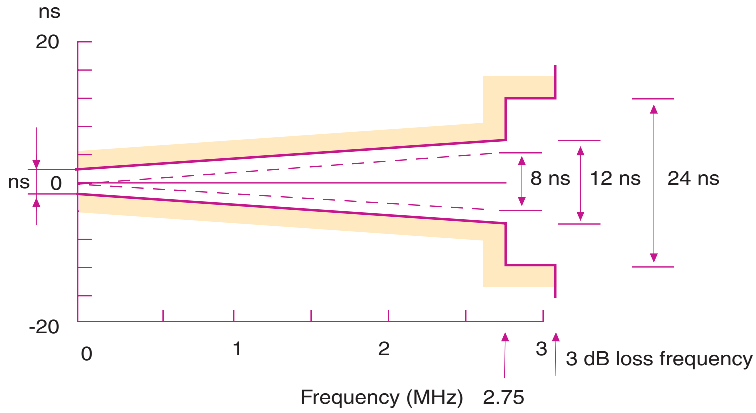
Specification for a colour difference signal filter used when sampling at 6.75 MHz



d)



e)



f)

d) Template for insertion los/frequency characteristics
e) Passband ripple tolerance
f) Passband group-delay tolerance
Note: The lowest indicated values in e) and f) are for 1 kHz (instead of 0 MHz)

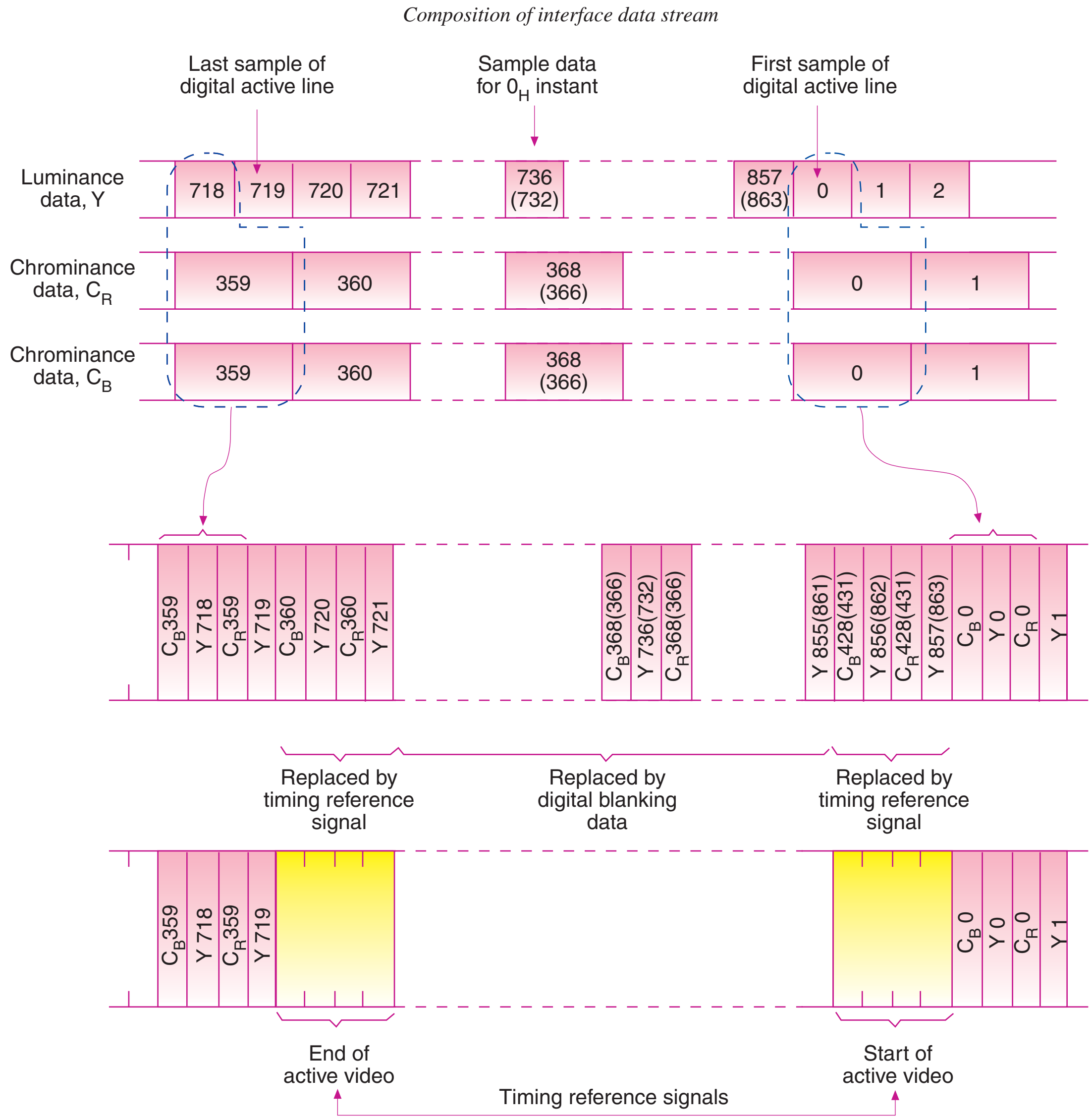


Figure 5 Composition of the data multiplex and position of the timing reference signals, EAV and SAV
Note: Sample identification numbers in parentheses are for 625-line systems where these differ from those for 525-line systems (see also Recommendation 803)

The meaning of an orthogonal sampling structure and co-sited samples are exemplified in Figure 2.

The encoding of all three signals is performed in A/D converters shown in principle in Figure 3.

The Low Pass Filters (LPF) in the figure are anti-aliasing filters to prevent the sidebands from the sampling process to overlap into the baseband signals. These filters have very tight tolerances regarding frequency and phase response (group delay). This is to ensure that the video signal can pass many A/D and D/A conversions without undue reduction in signal quality. The filter characteristics are shown in Figure 4. The same filters are also used as reconstruction filters in the D/A converters.

In order to interface the digitally encoded signals to external processing equipment, a standardized interface is needed. In Figure 3 this standardized signal structure is performed in the parallel data multiplexer. All details concerning this interface are described in ITU-R Rec.656-1, the interface standard [2], [3]. Of fundamental importance is the description on how the sample values from the co-site samples are organized in time. Figure 5 shows the composition of the data multiplex and the position of the Timing Reference Signals (TRS), EAV and SAV (End of Active Video / Start of Active Video).

With 720 active samples per line the duration of a digital active line is 53.3 µs, or 1.3 µs longer than the analogue active line. Figure 6 shows the relative positions of the digital and analogue lines.

It will also be seen that the analogue synchronizing pulse has disappeared in the digital signal and is replaced by the TRS

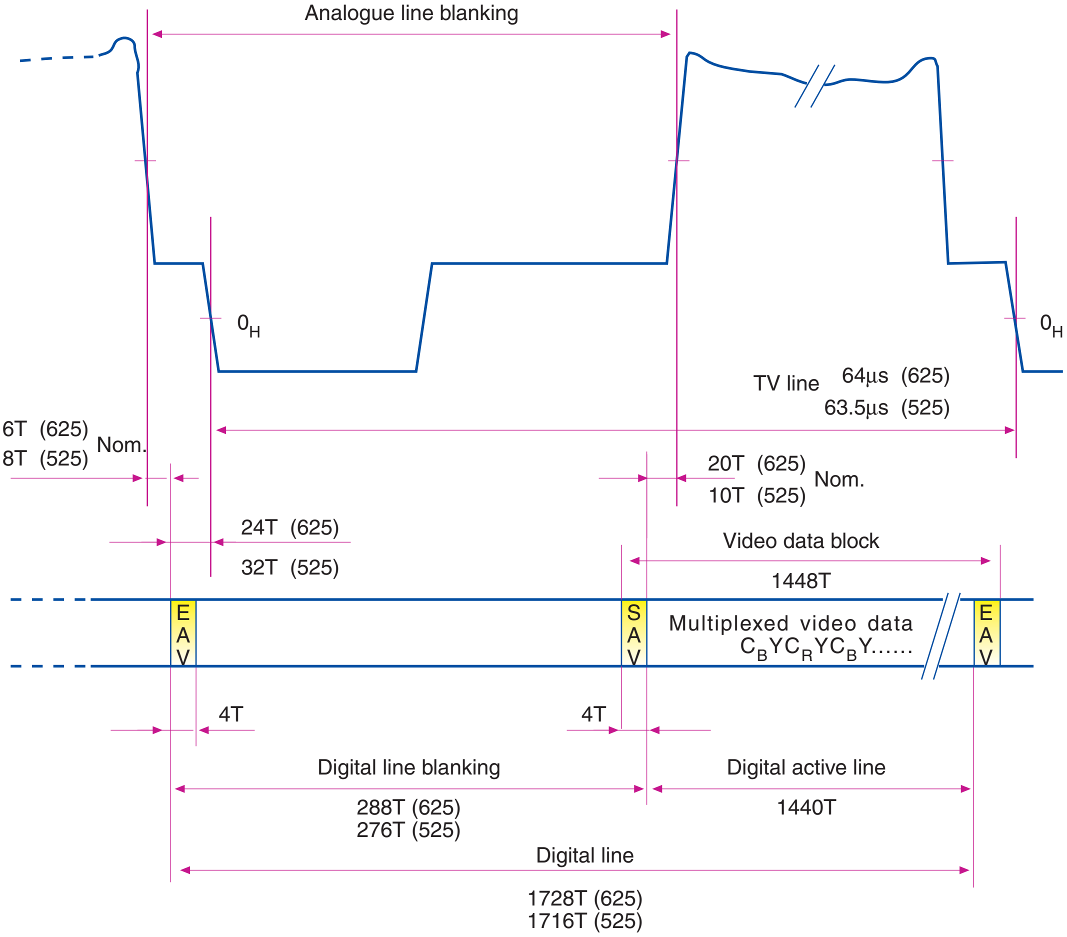


Figure 6 Relative positions of digital/analogue lines
T: clock period 37 ns nom.
SAV: start of active video timing reference code
EAV: end of active video timing reference code

words. These data words consist of a three word preamble and a fourth sync word. The preamble consists of the sequence *FF* 00 00 (hex). The TRS structure is shown in Table 2.

The three bits *F, V* and *H* carry the necessary information while the bits P_0 to P_3 are protection bits for the *F, V* and *H* bits. This fourth TRS word is *the only word* in the Rec. 656 data multiplex that is protected. Single error in the *F, V* or *H* bits can be corrected, dual errors can be detected. All other data in the multiplex are unprotected! This means that the quality of the channel transferring Rec. 656 signals must be very high.

The digital blanking interval between EAV and SAV constitutes what is termed an ancillary data channel. The horizontal blanking is used for transferring the audio that accompanies the video signal. The capacity is sufficiently high to carry a number of AES/EBU encoded audio signals. For the time being equipment for carrying four audio channels (two stereo pairs) have been developed. However, there is room for eight digital audio channels.

In addition to the bit-parallel interface, Rec. 656 describes also a bit-serial version. It is this serial interface that enables larger and practical digital television production systems to be designed and built. Figure 7 gives an overview of the serial interface. The parts that are standardized in Rec. 656 are the serial encoder and decoder.

The co-processors enable ancillary data like digital audio to be multiplexed into the video bit-stream and extracted at the receiving end. This possibility is a great advantage and can lead to substantial cost savings in larger installations. A detailed description of the components that make up the serial interface is outside the scope of this article. Further reading can be found in [4].

Neither the coding nor the interface recommendations bear any reference to the image format. Sampling frequencies and other coding parameters are only related to line frequency (line duration) and frame frequency. This means that when the wider 16:9 image format is introduced for 625 line television, digital pro-

Table 2 Structure of the TRS signal

Video timing reference codes[1]				
Data bit number	**First word (FF)**	**Second word (00)**	**Third word (00)**	**Fourth word (XY)**
9 (MSB)	1	0	0	1
8	1	0	0	F
7	1	0	0	V
6	1	0	0	H
5	1	0	0	P_3
4	1	0	0	P_2
3	1	0	0	P_1
2	1	0	0	P_0
1[2]	1	0	0	0
0	1	0	0	0

[1] The values shown are those recommended for 10-bit interfaces.

[2] For compatibility with existing 8-bit interfaces, the values of bits D_1 and D_0 are not defined.

F = 0 during field 1, 1 during field 2
V = 0 elsewhere, 1 during field blanking
H = 0 in SAV, 1 in EAV
P_0, P_1, P_2, P_3 = protection bits
MSB = most significant bit

duction systems based on Rec. 601 coding can be used with the same excellent picture quality as for the 4:3 image format.

Proposals have been made to increase the sampling frequencies in the same proportion as the picture width is increased, giving sampling frequencies of 18 MHz and 9 MHz for the luminance and colour difference signals, respectively. This proposal, should it be adopted, would have disrupting effects on the benefits of having a world standard for digital coding of video signals. In addition, the economical and practical repercussions would be considerable. It can also be argued whether the picture quality received in the homes would be visibly enhanced as a result of the increased sampling frequencies. Both transmission systems and receiver displays would have to be enhanced, something that is hardly economically feasible.

Digital video production equipment

Today all necessary system components that are needed to build a digital video production centre, have been developed and are available on the market. Until the end of 1992 there was much activity for digital composite TV production. But since early 1993 interests have been more concentrated around component digital systems based on the ITU-R Recommendations 601 and 656.

In the same way as modern analogue production systems could not exist without the VTR, so can the digital production system not exist without the digital VTR-(DVTR). This machine, with its ability to make an exact replica of the original digital signal, opens up possibilities which do not exist in analogue systems. But other system components are also important to (digital) television production, so let us have a look at some of them:

- *Vision mixer*
  This unit can cut, mix, fade and wipe between sources. In addition, high quality chroma key, mix/effects and multi-layering can be performed.
- *DVE*
  Digital Video Effect generators are normally used in conjunction with a vision mixer. Effects like picture shift, zoom, flip, spin, tumble, rotation, etc. (often with perspective) are common, and complicated key signals can be generated.
- *Caption generators*
  Caption generators provide text to be added to the picture. These units can generate all possible fonts and most of them provide anti-alias filtering for best picture quality.
- *Graphics*
  ‘Graphic units/paint boxes’ can generate diagrams, figures and drawings. Also a video frame can be grabbed in the unit’s memory and used as the basis for all kind of picture manipulation.
- *Still store*
  This is a unit for storing still pictures in a hard-disk based memory with short access time.
- *Synchronizer*
  When signals generated externally have to be processed, it is necessary to synchronize these signals to the local reference. A synchronizer loads the incoming signal in a picture memory and reads it out with a clock locked to studio reference. In this process the signal is delayed by (at least) one frame or 40 ms. If the accompanying audio is not multiplexed in the video bit stream through the synchronizer, the audio must be delayed separately by the same amount.

A common denominator for all these units are input/output interfaces according to Rec. 656. Together with picture sources like cameras and telecine-machines, complete studio/post production systems can be built; A/D converters and sync generators are necessary prerequisites. To be able to handle source signals from old PAL installations/recordings, digital decoders will also be needed.

When considering the system aspects for digital installations, a different set of problems become apparent compared to analogue systems. One property of digi-

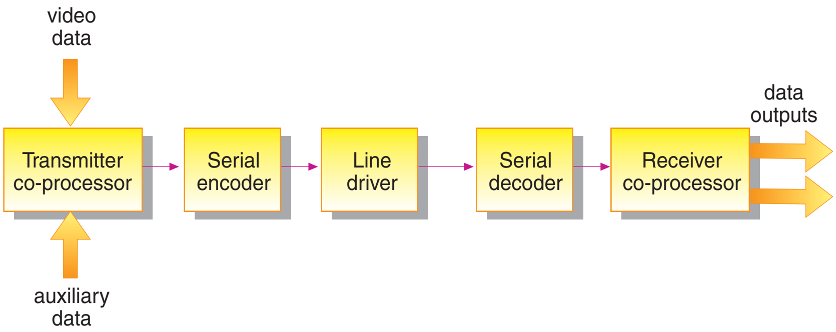


Figure 7 Serial interface overview

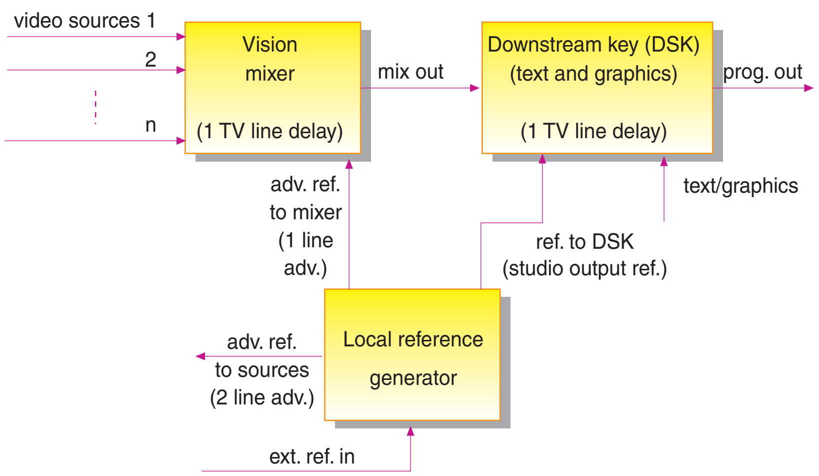


Figure 8 Reference distribution in a digital installation (in principle)

tal equipment is that processing takes longer time than with analogue equipment. Vision mixers normally have a processing delay of at least one TV line. The reason for this large delay is the need to synchronize all input sources to within the same picture sample. Since the mixer hardware performs parallel processing on multiplexed Y, C_B and C_B samples, all inputs must obviously be in step. If the mixer is large and complex, output synchronization may also be necessary adding one extra line of delay. Compared to the system reference signal, output video may be delayed two TV lines. For further downstream equipment running on the same system reference this delay will cause the output video to be shifted vertically, something that is unacceptable. To overcome this, it is necessary to have several reference sources, originating from the same generator but timed with respect to one another in such a way as to compensate the processing delays as indicated in Figure 8.

The last couple of years have seen the emergence of hard-disk based editing systems. These units are nothing more than a hard-disk recorder for digital video (and audio) with heavily bit compressed signals, giving a long playing time (several hours) on a stack of 1 Gbyte disks. With the short access time of a hard-disk, any scene in the source material can be viewed instantly. For editing purposes this is of great importance since edit time will be decreased and creativity increased. The hardware in these systems are standard 'boxes' as developed by the computer industry; only the software is specially developed for the purpose.

The short list above of digital video production equipment for broadcast use are normally found as dedicated products. The sale may be limited to a few hundred, maybe a few thousand units, giving the very high price normally found for this type of equipment. However, the performance and processing power of a standard PC has increased dramatically; the result is that a normal 486 PC with a high clock frequency can now process live digital video in real time. But still some concessions on picture quality must be accepted compared to 'Rec. 601 quality'. Nevertheless, it is only a matter of time before this hurdle is overcome. Already today 'top-rank' graphical systems and still stores run on standard computer industry platforms. This trend will open up a completely new situation for the broadcaster / program production companies with reduced investment costs and increased opportunities. And it will come sooner rather than later.

References

1 CCIR Recommendation 601. Düsseldorf, 1990.

2 *EBU interfaces for 625 line digital video signals at the 4:2:2 level of CCIR recommendation 601.* European Broadcasting Union Document, Tech. 3267E, 2nd ed. Geneva, 1992.

3 CCIR Recommendation 656. Düsseldorf, 1990.

4 Tancock, M. *Making the right connections: an overview of the Serial Digital Interface.* SONY Broadcast & Communications, Basingstoke, 1990.

Digital Television based on MPEG2 – Activity in Norway

BY GISLE BJØNTEGAARD

1 Introduction

There is an overall trend to represent information on digital form. This is also true for images – both still images and video. One of the requirements for practical use of digital video is digital image compression. This presentation will focus on why this is so, and why the transition to digital representation is taking place now. Another main part of the presentation is to outline the MPEG solution to digital video compression.

Section 2 will address the reason why compression is necessary and why it is so much in focus now. At the moment ISO/MPEG is the "main player of the game". Section 3 will therefore try to characterize MPEG as a group. Section 4 gives a general overview of "hybrid DCT" coding and the MPEG version of this type of coding in particular. Section 5 deals mainly with the NTR activity in developing the MPEG2 standard. Section 6 describes very briefly the implementation of MPEG2 at NTR.

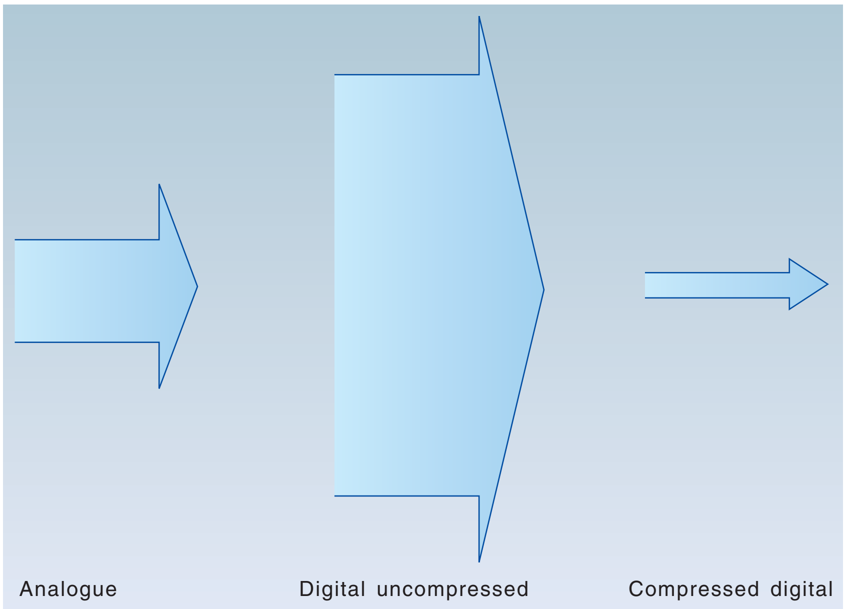


Figure 1 Illustration of bandwith requirements for transmitting video on analogue form, uncompressed digital and compressed digital

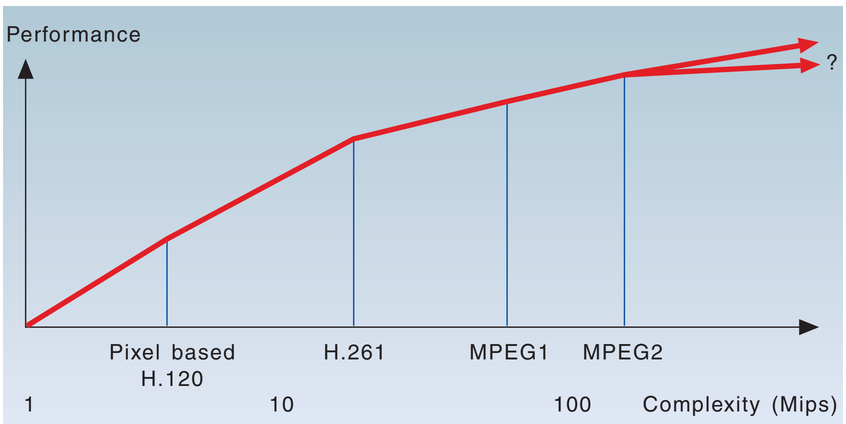


Figure 2 Relation between performance and hardware complexity of existing coding standards

2 Why is digital image compression so much in focus now?

2.1 Why is video compression needed?

When digitizing images, each point in the image is represented by a number. When the signal resulting from this process is being transmitted, it will require much more bandwidth than the original analogue signal. Transmission of the digital signal will typically require 4 times as much bandwidth as the analogue representation. We therefore say that the signal has been expanded by a factor 4.

After digital compression the signal is still represented on digital form, but the number of bits is reduced typically by a factor of 30.

The total result of the digitization and compression process is therefore that we have produced a digital signal that requires a transmission capacity of 10 – 20 % compared to the original analogue signal. The process is illustrated in Figure 1.

2.2 Connection between digital compression and available electronics

Digital image compression has evolved over the last twenty years. It basically started at the time when the first microprocessor was introduced. This linkage to the availability of suitable hardware should be no surprise since the image compression techniques are heavily dependent on computational power. The coding methods always have to adjust to the current status of implementability.

A live video sequence typically consists of 1 – 20 Mpixels/sec depending on the quality. The simplest compression methods operating on a pixel to pixel basis requires a few instructions per pixel (Ipp). Used together with a video format of 1 Mpixel/sec this results in a hardware requirement of a few Mips. This turned out to be a demanding task ten years ago – especially since the types of operations needed are very "data dependent" and therefore cannot be "hardwired".

About 10 – 15 years ago the trend was to move from the pixel to pixel coding methods to transform or vector quantization based coding methods. These methods are considerably more powerful concerning compression. The penalty is that they also require more computation power – typically 10 – 20 Ipp.

The next large step in both performance and required processing power was the introduction of motion search and motion compensated prediction. This increased the requirement for processing power by almost a factor of 10. The standards H.261 for video telephony and MPEG1 falls into this category (1,2,3).

With MPEG2 (4) the complexity of the coding method was increased yet another step by allowing more prediction modes and introducing B-frames (used also in MPEG1, see definition below).

This development has been tried visualized in Figure 2. There is no reason why this development should stop now. We will certainly see new more powerful/more complex coding methods being developed hand in hand with the increased availability of computing power. One interesting trend is that at the moment compression performance does not seem to increase very much as complexity is increased. Is this a "local" phenomenon, or will we see a saturation point regarding compression?

3 MPEG2

3.1 The mandate of MPEG

MPEG (Moving Pictures Experts Group) is the nickname of a working group in ISO. Its full name is ISO/IEC JTC1/SC29/WG11. Originally the mandate of MPEG was to develop standards for storage of video. Concerning standardization of real time transmission systems like video conferencing and television, this has been the responsibility of telecommunications bodies like CCITT/CCIR (now ITU).

So why is MPEG so heavily involved in standardization for digital TV? There is definitely competition between international standardization groups concerning where particular issues belong. In the case of MPEG the trick for including other applications than storage was to redefine the target to produce "generic" standards for audio-visual services. The meaning of "generic" here is to develop core compression techniques that may be fitted into a variety of different applications. This means that the area of applications is almost unlimited. MPEG has furthermore turned out to be a powerful group for several reasons:

- MPEG comes at the right time (referring to the above mentioning of relation to available computational power).
- MPEG consists of participants from the main electronics industry who are interested in having a standard now.
- The work method of MPEG is suitable because development is performed inside the group and because the work is done fast!

MPEG2 is dealing only with the "core" techniques for video and audio compression in addition to multiplexing. Other issues like bit error handling and modulation will be service dependent and will have to be standardized for each service area.

3.2 What happened to the other standardization bodies?

The ITU body responsible for standardization of video-conferencing with high quality decided to work jointly with MPEG to develop a common standard. By this co-operation the ITU group had the opportunity to influence the standard in important areas for real time communications. The main issues specific for real time communications were:

- To ensure low coding/decoding delay
- To take due care of bit error statistics in the ATM network
- To ensure compatibility with existing standards for real time communication – in particular H.261.

The international body responsible for standards in the TV area is another part of ITU. There has been some effort from this side to initiate work on digital TV. However, this body has not succeeded in establishing any significant activity. The fact is rather that interested participants have concentrated on the work going on in MPEG even if it is outside the mandate of that group to standardize digital TV. It is foreseen that ITU will take the outcome of MPEG2 and make the necessary extension concerning service aspects, etc. to fulfil a complete TV standard.

In Europe, EBU has traditionally been responsible for making standards for TV broadcasting. EBU has not been directly involved in developing the core techniques for source coding of video and audio. On the other hand, EBU has been active in the area of modulation, etc.

A new project – DVB (Digital Video Broadcasting) – has been started in Europe. The mandate is to make complete European standards for digital TV based on satellite, cable and terrestrial transmission media.

Figure 3 illustrates the position of MPEG in developing core technologies used by other bodies to define application standards.

4 Main characteristics of the MPEG coding method

This section gives an overview of the main aspects of the MPEG coding method.

4.1 Pixel representation and block division

Each picture is represented by a number of horizontal and vertical points or pixels. For a standard TV picture the number of pixels in a picture is 720 x 576. Each pixel is represented by 8 bits of 256 grey levels.

The collection of pixels is divided into blocks of data. In MPEG "block sizes" of 16 x 16 (Macro Block) and 8 x 8 (Block) are used. These block sizes are the units used during the coding process.

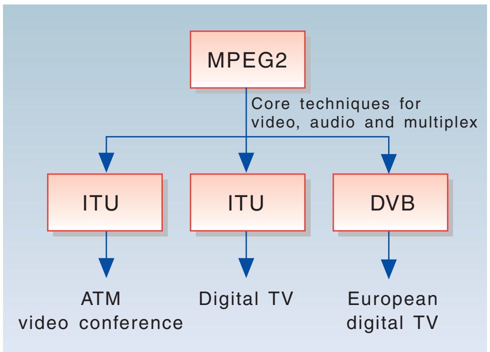


Figure 3 Relation between MPEG2 and the other bodies concerned with video standards

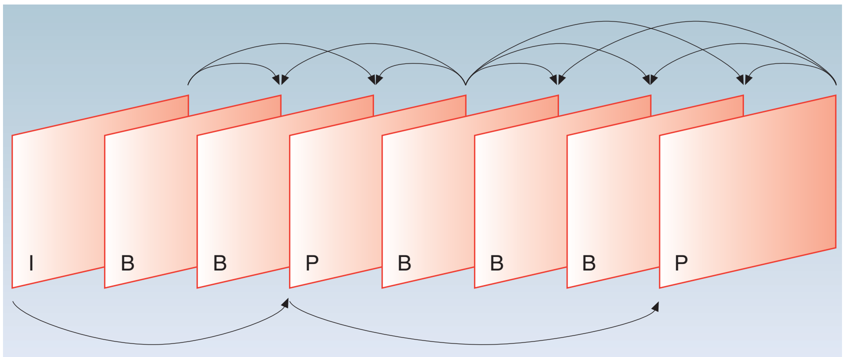


Figure 4 Picture types in MPEG

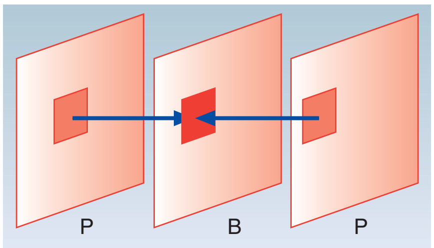


Figure 5 Different prediction modes for B-pictures

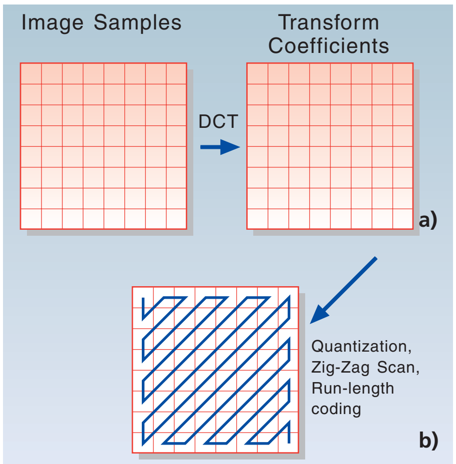


Figure 6 a) 8 x 8 Discrete Cosine Transform (DCT); b) zig-zag scanning

4.2 Prediction

A major source of compression is prediction. The prediction in MPEG is block based, which means that the block to be coded is predicted from parts of the video sequence already coded. Assume that the block to be coded is:

$O(i,j) \qquad i,j = 0..15$

Assume that we can find a prediction:

$P(i,j) \qquad i,j = 0..15$

This prediction must be available at both the encoder and decoder end. We therefore have to send the difference *D()* of the two blocks to the decoder:

$D(i,j) = O(i,j) - P(i,j) \qquad i,j = 0..15$

The block *D()* normally contains much less information than the original block *O()*.

At this point it is necessary to introduce the three picture types defined in MPEG (see Figure 4):

- I-pictures are coded without prediction – or *P()* is set equal to 0. The I-pictures serve as "entry points" to the video signal since no pervious information is needed to reconstruct a picture.
- P-pictures are predicted from the last P-picture by motion compensated prediction. This means that horizontal and vertical "vectors" define *P()* as a displaced block from the last P-picture.
- B-pictures. This picture type is optional. If present, B-pictures require that both previous and forward I- of P-picture is available. The prediction *P()* may be forward, backward or a combination of the two as indicated in Figure 5.

4.3 Transformation and quantization

The block *D(i,j)* is divided into four 8 x 8 blocks and a two dimensional Discrete Cosine Transform is performed (see Figure 6a). This operation has the effect of shifting most of the energy of the signal to a few transform coefficients.

The transform coefficients are quantized. This means that the coefficients are represented by a number of levels. The spacing of these levels are very important. Increased spacing will reduce the information in the signal. The result is that less capacity is needed to represent the signal and also that the reconstructed signal deviates more from the original.

4.4 Scanning, coding and inverse transform

The quantized transform coefficients are next "zig-zag" scanned (see Figure 6b) and coded by a two dimensional Huffman code. A bitstream is then constructed.

To reconstruct the deviation from the prediction, an inverse DCT is performed. Due to the quantization process, the reconstructed difference differs from *D()*:

- $D'(i,j) \neq D(i,j)$
- $D'(i,j) - D(i,j)$ is a measure of the error in the reconstructed picture.

The above description represents MPEG1 coding. At the same time it shows the elements in a "hybrid DCT" coding scheme. By this we mean the mixture of prediction and DCT of the deviation from the prediction.

The main addition in MPEG2 is that provision is made for coding interlaced pictures. This means that a picture or a frame is a combination of two fields which represent two positions in time. This is shown in Figure 7. A block here consists of a mixture of the two fields. To cope with this kind of picture, some additional prediction modes are introduced. In addition the coding may be made "field oriented" or "frame oriented". The details will not be given here, but some typical aspects will be treated in the next section.

5 Activity at Norwegian Telecom Research

5.1 Proposal for the initial MPEG2 test

The MPEG2 work was started by invitation for proposals to coding methods. The rules for the proposals was set up in early '91. Presentation and testing of the proposals was performed in November

’91. The proposals should include the following items:

- A complete documentation of the coding method
- An evaluation of implementation complexity of the method
- Coded and decoded sequences of predefined picture material. Four different test sequences, each of 5 seconds’ duration were coded. The coding bitrates were specified to be 4 and 9 Mbits/second.

Altogether 30 proposals were submitted for this initial test. The actual subjective test of all the sequences was performed in parallel with an ordinary MPEG2 meeting and took 3 – 4 days.

In addition to the picture quality test, there was an assessment of implementation complexity. A group of 5 – 10 hardware experts made a ranking of the 30 proposals.

The NTR proposal came out among the 3 best concerning image quality as well as implementation complexity. No other proposal showed similar results both in subjective quality and implementability.

5.2 Inputs to the MPEG work

After this initial subjective test, a “Test model” (TM) was established for further development. The name “test model” had nothing to do with the initial test. The purpose of this TM was to establish a model that would serve as a reference during the improvement of the coding method. The first TM was established in early ’92 and it was improved until March ’93. At that time the version of the TM (TM5) was taken as the basis for the final definition of the standard.

Several proposals for improving the test model were made by NTR. They were all focused on giving a good performance/implementability relation. Two of the main proposals which were included in the TM will be covered below. They are both concerned with improvements of the prediction part of the coding process. Prediction is probably the most important area for improving coding efficiency. The following two sections go right into the base elements of making a good coding method. This also gives a good illustration of the kind of problems that have to be addressed in developing a compression standard.

5.2.1 Motion estimation

A major part of the coding process is to predict a 16 x 16 block of image data from a set of displaced blocks from previously coded pictures. The simplest case with prediction from one previously coded picture is illustrated in Figure 8a. *New* represents the block to be coded. *Old* represents a displaced block from the previous picture. *V* is the motion vector.

The accuracy of *V* is 1/2 pixel both horizontally and vertically. The meaning of this is shown in Figure 8b. Here we see 4 possible predictions of the pixel *p*. The pixels where the prediction is taken from are marked *a,b,c* and *d*. The 4 predictors shown in Figure 8b are:

a	Integer pixel displacement both horizontally and vertically
(a+b)/2	Half pixel displacement horizontally, integer pixel displacement vertically
(a+c)/2	Integer pixel displacement horizontally, half pixel displacement vertically
(a+b+c+d)/4	Half pixel displacement both horizontally and vertically.

The half pixel accuracy in the motion vectors serve two purposes:

- The definition of displacement is more accurate with half pixel resolution than by integer pixel resolution.
- The different predictors above represent different filters in the prediction process.

The last point turns out to be the most important one. The motion vector may be considered to consist of two parts: The integer part which gives the motion and the remaining half pixel part which gives the filtering.

The motion vectors are obtained at the encoding side by “block match” operations. Taking the above considerations into account, it turns out that the motion search works best if performed in the following two steps:

- Do integer pixel search where the block to be coded is compared with the original previous picture. This gives the best estimate for the motion *V*0.
- Do half pixel search in 9 positions around *V*0 as shown in Figure 8c. This search is performed in the previously decoded picture to obtain optimal filtering in the prediction loop.

From the beginning both integer and half pixel search was made on original images. The NTR proposal for doing integer pixel search in the original picture and half pixel search in decoded picture reduced the overall bitrate by about 10 %.

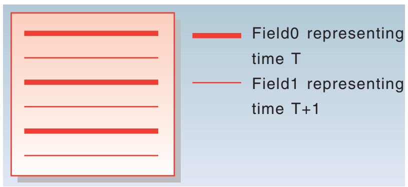


Figure 7 A picture or frame consists of two fields

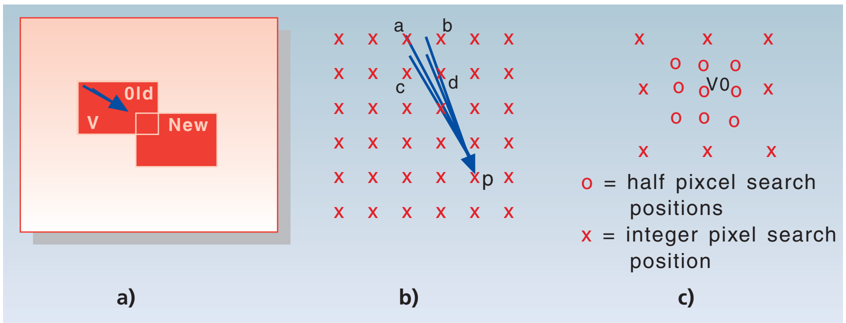


Figure 8 Illustration of motion compensated prediction. a) Prediction by displaced block from the previous picture; b) half pixel motion vectors; c) position of integer and half pixel positions in the motion search

5.2.2 Dual´ prediction

It had been demonstrated that the quality of prediction improved significantly if averaging the prediction from two fields or pictures was allowed. This is one of the mechanisms used in B-frame prediction (where "B" stands for bi-directional). The essential point with "B-frame coding" is illustrated in Figure 9a. The coding order is: picture 4 followed by pictures 2 and 3. Pictures 2 and 3 may then be predicted from both pictures 1 and 4 and this bidirectional prediction turns out to be very efficient for certain types of pictures.

A similar benefit may be obtained without reordering the sequence of coding pictures. One version of such a predictor has been named "dual´". The essence of this prediction mode is shown in Figure 9b. The macroblock to be coded consists of two fields. Basically one motion vector *V* is used to define the prediction of both fields. The prediction of each field is made up of the average of predictions from the two fields of the last coded picture. The displacements of the different parts of the predictions are obtained by linear scaling of the vector *V*.

This prediction mode was first proposed by NTR and included in the standard. Use of this prediction reduces the bitrate by 10 – 20 % for pictures with moderate motion.

6 Hardware implementation

It is important to point out that MPEG2 is an interface standard rather than a standard to specify how to build equipment. Only the syntax is specified so that bits generated from different equipments may be decoded correctly. MPEG2 may also be said to be the decoding standard specifying the interface to be used on the encoder side. There are therefore numerous possibilities of making implementations that will satisfy MPEG2. This also means that detailed knowledge of the coding mechanisms is needed to make good encoder implementations.

At NTR we are making an MPEG2 encoder implementation. We have a tradition of making equipment with favourable performance/cost relation. This is also the approach this time. The main characteristics of the encoder may be summarized:

- Main profile at main level (MP@ML)
- Use of two proprietary ASICs for motion search and prediction
- Use of one proprietary ASIC for bit stream generation.

The encoder will fit on one board and is planned to be ready early 1995.

7 Conclusion

The MPEG standard defines how to do digital compression of video so that the required transmission capacity is only 10 – 20 % compared to analogue TV. This means that TV programmes may be brought out to the user much cheaper than before. An obvious result of this is that we will see more TV programmes brought out to the user in the future.

Another aspect is the digitization of the video signal. This means that it will be more easily connected with other data services.

These two aspects: digital video which is more available than before will most certainly change the use of video in the future. The technical elements are now ready. The next step will be to define applications which make use of the technical possibilities.

Norwegian Telecom Research have been very active in developing most of the present video compression standards. Concerning MPEG2 we have taken an active part in the development of the core coding elements.

As a result of this, we are now undertaking hardware implementation of the full MPEG2 video algorithm.

8 References

1 CCITT. *Video Codec for Audiovisual Services Recommendation.* Geneve. (CCITT H.261.)

2 ISO. *Information technology – Coding of moving pictures and associated audio information for digital storage media at up to about 1.5 Mbit/s.* Geneve 1993. (ISO/IEC 11172-2.)

3 Le Gall, D. MPEG: a video compression standard for multimedia applications. *Communications of the ACM,* April 1991.

4 ISO. *Information technology – Generic coding of moving pictures and associated audio information.* Geneve 1993. (ISO/IEC CD 13818-2.)

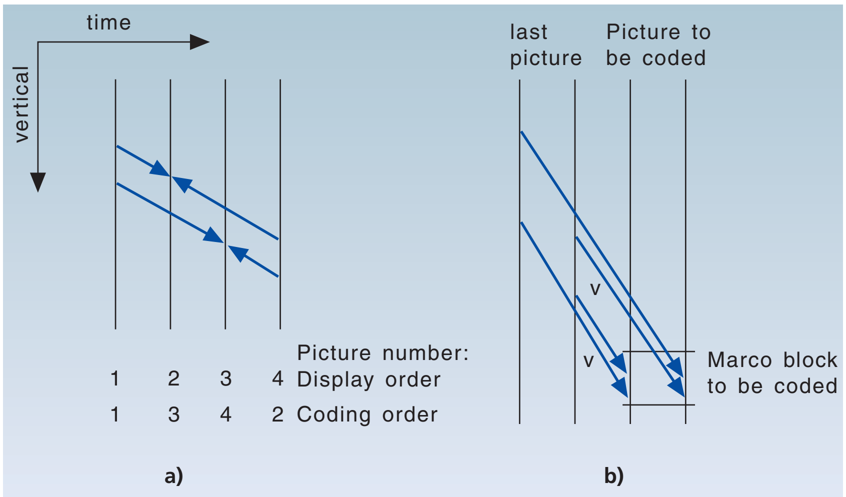


Figure 9 a) Use of B-frames; b) dual´ prediction

HD-DIVINE – a Scandinavian Digital Television System

BY PER APPELQUIST

In the summer of 1992 a digital terrestrial HDTV broadcasting system was demonstrated as a result of a collaborative Scandinavian study. The system included a motion compensated hybrid DCT video codec, a 512 carrier OFDM 16 QAM modem and four ISO/IEC Layer II sound codecs. The complete system was implemented in hardware. Since then further refinements of the vision codec and final assembly of the modem have resulted in a fully operational demonstration system. The current implementation uses the 1250/50/2:1 studio standard and fills an 8 MHz UHF channel. A second implementation allows reconfiguration from 1xHDTV to 4xCCIR Rec. 601 input signals.

1 Introduction

HDTV has been a top issue in discussions of the development of television since the early eighties. Still, there is just one regular service in the world today broadcasting in the HDTV format. The number of households that have access to this HDTV service is limited. With this in mind one has to ask why? There can be several possible explanations. One is that HDTV is still too expensive for the households as well as the TV networks. Another explanation is that the technology developed for HDTV did not have the potential to deliver a practical solution. Digital technology has meant a major breakthrough for broadcasting. What was impossible just a few years ago is now viable with digital solutions. Maybe the most important benefit with digital lies in the fact that it is so widely exploited by major industries such as telecommunication and computer industries and the synergies this might create in the future. The most obvious synergy is of course that the industries can benefit from the joint technology that now could be developed with the combined efforts. Another major synergy that has just recently become visible is that the services and products now produced by each industry separately can and probably will merge. This will create a major future industry of communicating and processing information and entertainment. Perhaps the digital technology will turn HDTV into a success.

The major obstacle down the road is the necessity to reach common standards. Standardization is a key activity in order to eliminate friction in the communication highways of tomorrow. On the other hand to establish standards means to freeze the technological development for a while. It must thus be timed very carefully before several de-facto standards are already established but not until the technology is reasonably mature.

In early 1986 the coding of standard television signals down to 34 Mbit/s was not considered viable to give broadcast quality pictures. Later that year the Swedish Telecom proved the opposite within the COST 211 bis programme by demonstrating a component digital video codec at 34 Mbit/s that not only delivered broadcast quality pictures but studio quality pictures. In 1992 it was believed in Europe that HDTV could not be compressed down to a level to fit into a terrestrial channel but in the summer that year a Scandinavian consortium proved at the International Broadcasting Convention at Amsterdam that this was in fact possible. This shows that the state of the art in digital video compression technology has been a moving target during the last 10 years. Perhaps the discipline is mature for standardization today.

HD-DIVINE was developed to show the feasibility of digital technology for television broadcasting. For this reason the combination of HDTV and terrestrial was made. Now this goal is achieved. For real systems to be launched for broadcasting services the concept must be varied. For instance the use of satellites must be taken into consideration as well as the delivery of standard quality television signals.

In the following the HD-DIVINE system is presented in technical terms.

2 The HD-DIVINE system

HD-DIVINE is a project that developed a demonstration system to prove the feasibility of HDTV in terrestrial channels. The system as it stands today will not be proposed as a standard neither in Europe nor anywhere else. It comprises a video codec and an OFDM modem as the main building blocks.

Since the first demonstration at the International Broadcasting Convention in July 1992 a demonstration system of a modern teletext service has been developed. It was felt that this area of television was not fully covered in the discussion of the next generation of television broadcasting standards. This Broadcast Multimedia system was demonstrated at the International Television Symposium in Montreux in June earlier this year.

A second codec was also developed with minor changes compared to the first. The most important added feature is the possibility to split up the service into 4 parallel television services each with 4:2:2 subjective video quality and stereo sound. The future plans include a further updated MPEG2 based video coding algorithm, a more elaborate service multiplex, a Broadcast Multimedia function and further studies of the basic modulation parameters for the terrestrial modem in connection with field trials.

2.1 The video coding algorithm

There are many vital building blocks in a terrestrial digital HDTV system, one of which is the video coding algorithm. The algorithm used in HD-DIVINE should demonstrate that acceptable image quality is achievable in the narrow terrestrial channel.

Early simulations indicated that the available bitrate on the channel is about 27 Mbit/s. In HD-DIVINE 4 x 128 kbit/s are set aside for audio, 64 kbit/s for data and 2 Mbit/s for the video data. The four mono audio channels are coded according to ISO/IEC layer II.

The source video signal has 1250/50 lines/frame and 25 frames/s. This signal is sampled at 54 MHz for the luminance and 13.5 MHz for the chrominance. This gives a total bitrate for the video signal of 648 Mbit/s if 8 bits/ pixel is used. To squeeze this signal down to 24 Mbit/s, which is the net datarate for the video part in HD-DIVINE, is a formidable task.

The algorithm used in the HD-DIVINE project is of the hybrid-DCT type. As the name "hybrid-DCT" implies, the algorithm used is a combination of different methods of which the most important are DPCM (Differential Pulse Code Modulation) and transform coding.

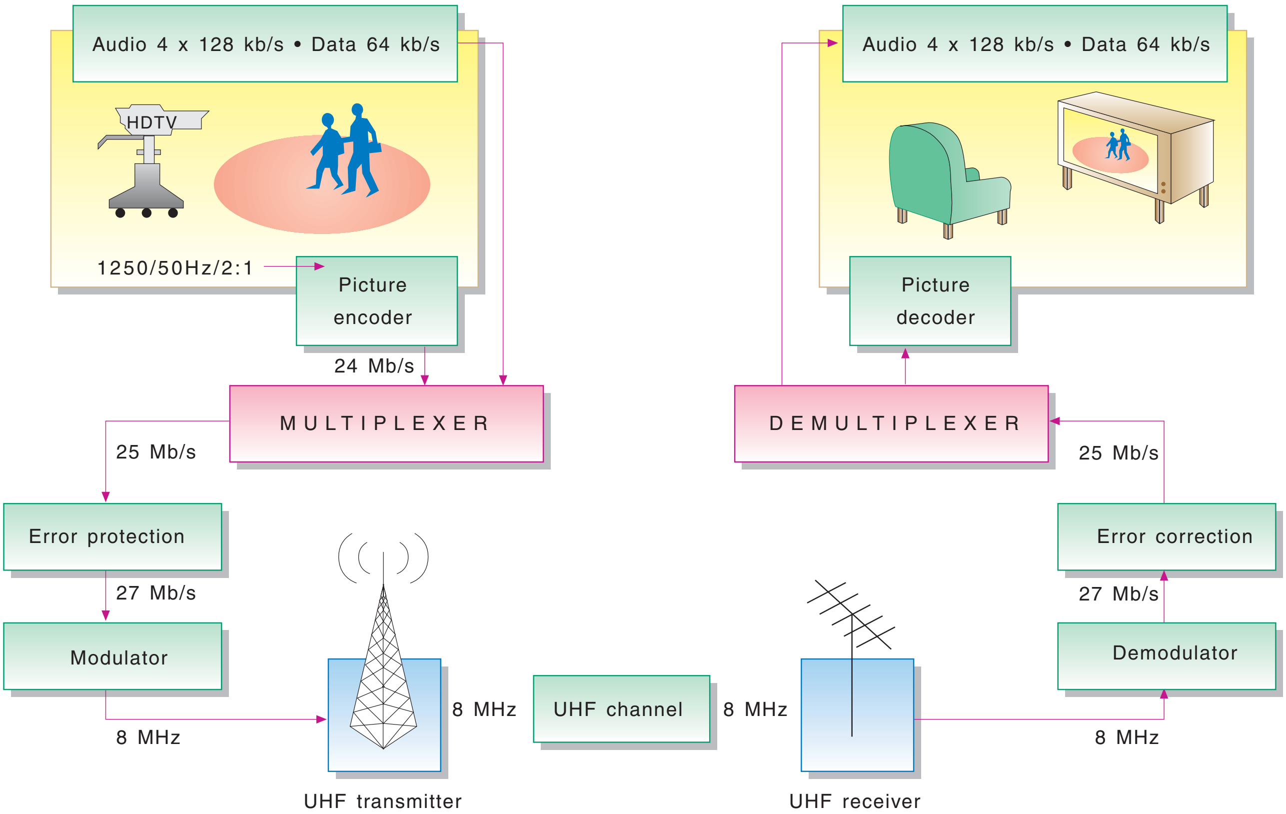


Figure 1 The HD-DIVINE system

2.2 DPCM

When using DPCM, a difference signal is sent over the channel. This difference signal is formed by subtracting a prediction from the original signal. A good prediction will give a difference signal (also called "error signal" or "residual signal") with a very skewed distribution. A perfect predictor would give a difference signal that is equal to zero, whereas a real predictor yields a signal which is not equal to zero. A usable predictor will, however, give a residual signal that can be compressed by subjecting it to variable-length coding.

Fixed-length coding of values from a signal with a skewed distribution is not the best (most compact) representation. A more efficient method is to give short code words to values of high probability and long code words to values of low probability; this is known as "variable-length coding". There are a number of methods that construct code hooks according to this principle, and the method used by HD-DIVINE (and most other video coding schemes) is Huffman coding.

A good predictor for moving images can use the previous image in the sequence as basis for the prediction. As the greater part of an image in a sequence does not change from one frame to the next, the residual image will be mainly zero or close to zero. This results in a skewed distribution of the residual image, which can be efficiently compressed by variable-length coding. By using movement-compensated prediction the result of the prediction can be even better. This is done by first estimating the apparent movement of each part of the image, between successive images. The predictor will then use this information to make a better prediction for the next image in the sequence. The coder function that does this is known as "motion vector estimation". It must be remembered that there is a cost associated with this type of prediction. The movement vectors must be sent over the channel to the decoder. The algorithm must be designed so that the gain from coding the residual image is bigger than the loss due to sending the movement vectors. The vital parameters are the number of movement vectors, the magnitudes of the movement vectors and the quality of the movement vector field.

2.3 Transform coding

If the DPCM part in a hybrid-DCT coder reduces mainly the temporal redundancy of the moving sequence, the transform-coding part is aimed at reducing the spatial redundancy and exploiting the characteristics of the human visual system.

Images are usually easier to describe in the transform domain than in the time domain. This is due to the fact that the DCT concentrates the energy into a small number of coefficients. For most images, the DCT is very close to the optimal transform (Kahrunen-Loeve) in this respect, but it is much easier to compute. HD-DIVINE uses an 8 x 8 pixel transform and, for many blocks, just a few (from two to four) coefficients are transmitted out of the total 64.

Another advantage of a DCT coder is that it is easy to include the frequency response of the human visual system. This is done in the quantization step. Every coefficient in the transform domain represents a spatial frequency (not exactly but closely enough), and is quantized according to the frequency response of the human visual system. Typically, high spatial frequencies are quantized with larger steps than low frequencies.

Furthermore, the thresholding of the quantized coefficients results in local adaptation according to the image content. The combination of the concentration of energy into a small number of coefficients, the quanization of the coefficients and the thresholding of the quantized coefficients results in a signal which is well-suited to variable-length coding.

The quantizing operation is also used to control the resulting bitrate of the coder. This is done by controlling the level of quantization according to the degree of output buffer filling.

2.4 Combination

HD-DIVINE makes the standard combination of DCT and DPCM, i.e. it applies a discrete cosine transform, quantization and variable-length coding to the output of a DPCM loop (Figures 2 and 3). It is important that all predictions are made on the basis of previously transmitted information. If this were not the case, the coder and decoder would not be able to make matching predictions and this is central to the operation of a hybrid-DCT system.

The decoder receives the residual image and movement vectors which are to be used in the prediction. The movement vectors are used in the decoder to movement-compensate the decoded image and this predicted image is then added to the residual image after variable-length decoding, inverse quantization and inverse transformation. This sum forms the output of the decoder.

There is one problem in this scheme: how is the process started? The first image after start-up cannot be predicted from a previous image, so this image must be treated in a special way known as “intracoding”. Intracoding does not use the predictor at all; the image is coded as if it were a still image, using only the transform coding part of the hybrid-DCT coder. This is also done at regular intervals to prevent the propagation of errors in the decoder. The use of intracoding at regular intervals is also known as “refresh” and is usually done a few times (once to five times) every second. The normal mode of the coder, using the predictor and transmitting the residual image together with the movement information, is known as “intercoding”. Usually the number of bits used for an intracoded image is much higher than the number of bits used for an intercoded image. There are rare cases where it is more efficient to transmit an intracoded image, typically when there has been a scene change or when the movement estimator has been unsuccessful in finding good movement vectors. The coder has the ability to choose locally between inter-/intracoding and it always uses the most efficient method.

How does HD-DIVINE differ from other hybrid-DCT coders? The most important differences are:

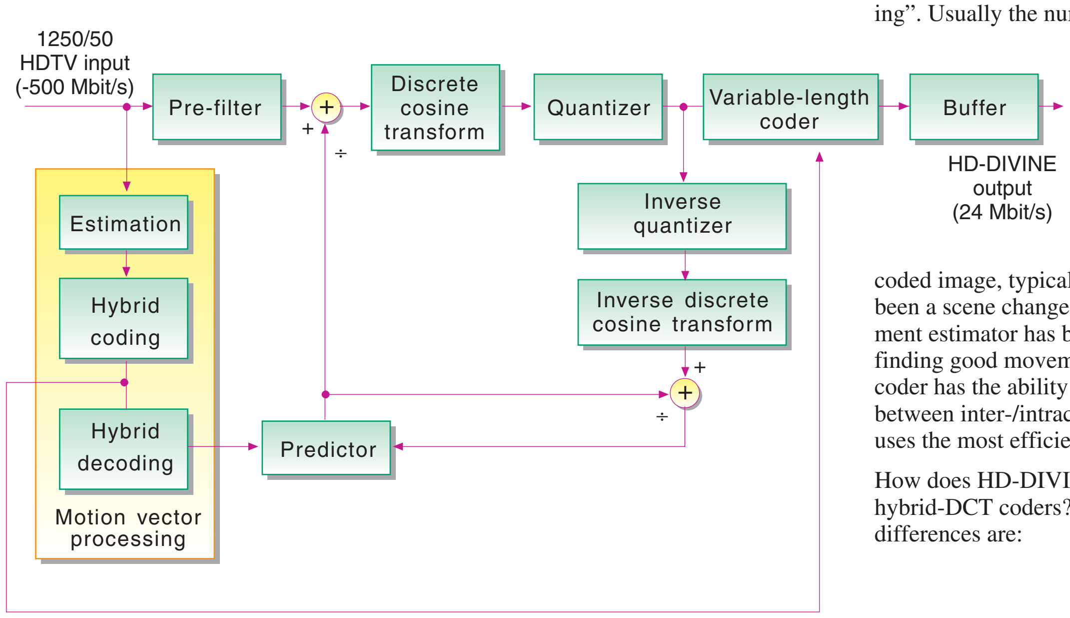


Figure 2 HD-DIVINE Hybrid encoder

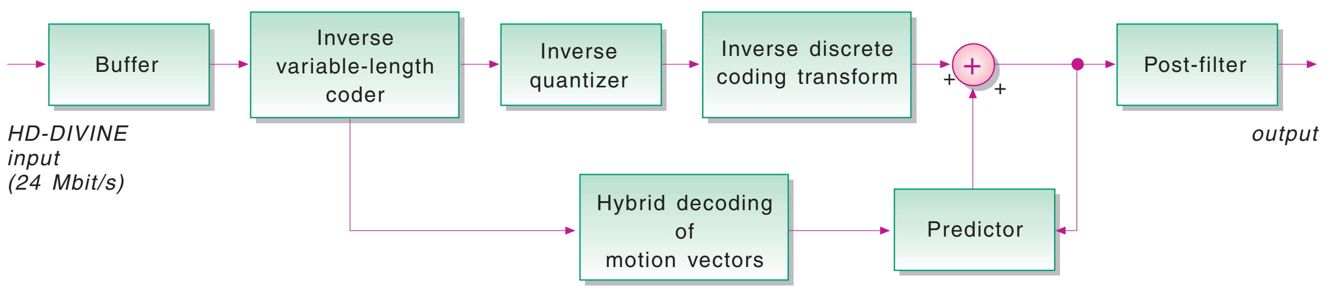


Figure 3 HD-DIVINE decoder

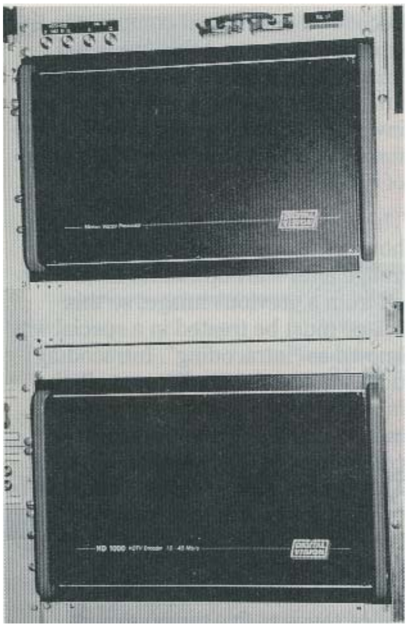


Figure 4 HD-DIVINE motion estimator and encoder

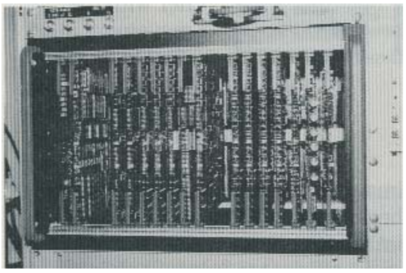


Figure 5 HD-DIVINE decoder

Figure 6 HD-DIVINE modulator

- The format of the movement vector field
- The coding of the movement vector field
- Adaptive pre- and post-filtering of the signal.

As can be understood from the above description of a hybrid-DCT coder, the quality of the predicted image is of vital importance. The better the prediction image the smaller the number of bits expended on the residual image. However, there is a cost associated with sending the movement vectors to the decoder. It is a delicate task to balance the gain in coding the residual image against the cost of coding the movement field. The trend in modern image coding is to spend more bits on prediction and less on the residual image.

The magnitude of the movement vectors is also important. There are extremely large excursions between successive images in a video signal. If the quality of the prediction is to be high, it is vital that the movement field can describe such large excursions. But, again, there is a trade-off between the magnitude of the vectors and the coding gain. Large vectors require the transmission of more bits to the decoder.

A very dense movement field is used in HD-DIVINE. There is a movement vector for every 4 x 2 pixel block in the image. This makes very precise prediction possible compared to older systems that use only a single movement vector for an 8 x 8 pixel block or for a 32 x 16 pixel block. Furthermore, the magnitude of the movement vectors is large, ± 32 pixels (horizontal) and ± 16 pixels (vertical). Furthermore, the movement estimation algorithm used gives half-pixel accuracy. Estimation and compensation is done on a field basis and is done only between fields of the same parity. The type of movement vector field used in HD-DIVINE needs efficient coding in order to leave some bits for the residual image. Due to the movement estimation algorithm used, the movement field is very precise, it describes the movement correctly and is noise free. One possibility for coding the movement field is to use DPCM, but the high quality of the movement fields is such that it is more suitable for hybrid-DCT. Accordingly, the movement vectors are subjected to the same treatment as the image data, in a hybrid-DCT loop, but with one exception: there is no movement compensation of the movement field. During this coding some minor distortion of the movement field occurs, but, in spite of this, the quality of the prediction remains very high.

It is not always possible to reduce the amount of data in a video sequence by 95% without introducing visible distortion. Some scenes will have perceptible distortion and artefacts. When the images in a sequence are difficult to code the quantization of high-frequency components is very coarse. This may result in visible ringing around edges and the image may also contain noise and artefacts (such as blocking effects). One way to make images easier to code is to lower their spatial-frequency content. This results in fuzziness. However, it may be a good idea to exchange the type of distortion introduced by the hybrid coder with fuzziness. An adaptive pre-filtering process analyses the images and, in areas where the hybrid-DCT coder risks introducing distortion, such filtering reduces the resolution marginally.

Adaptive pre- and post-filtering has another very significant effect on image quality. The filters used are very efficient in removing noise. In the pre-filter this is used to reduce source noise. Noise is not only unpleasant to look at: it also costs bits to code. Consequently, the noise-reduction properties of the pre-filter are doubly important. The post-filter reduces the quantizing noise that is introduced in the coding loop. The post-filter can also perform aperture correction, but this has not yet been tested.

2.5 Modulation

Perhaps the most interesting part of the HD-DIVINE project is the modulation scheme. The technique is not new and is currently also used in digital audio broadcasting (DAB). The actual implementation, however, differs in the choice of modulation parameters.

The terrestrial channel is a difficult one, particularly due to multipath propagation. It can be questioned whether a single-carrier system can be robust enough in such an environment. One way to obtain the desired capacity in the channel is to use coded orthogonal frequency division multiplexing (COFDM). COFDM is a multi-carrier system and it combines modern forward-error protection algorithms with very efficient implementation.

In a COFDM system a large number of low symbol-rate carriers are used, as opposed to single-carrier systems with one high symbol-rate carrier. HD-DIVINE is a 512-carrier system in which only 448 carriers are actually used. The reason for not using all 512 carriers is the need to shape the spectrum of the COFDM signal to match the PAL spectrum and to eliminate adjacent-channel interference. In an environment where PAL and COFDM have to coexist it is desirable to minimize the interference between the two types of signals. The main problem is the large image and sound carriers in PAL. The solution is to open up slots in the HD-DIVINE spectrum and thus prevent destructive interference. Most of the 64 unused carriers (about 40) form the slots in the COFDM spectrum. Some are left unused on each flank of the spectrum, due to adjacent-channel interference (Figure 7). The experimental modem is programmable in respect of the number of unused carriers and their positions in the channel.

Quite apart from the slots and flanks mentioned above, the COFDM spectrum is optimally shaped. The spectrum of each closely-spaced carrier overlaps that of its neighbours, which results in a very flat spectrum for the total signal and an efficient use of the available bandwidth.

Demodulation of the carriers is possible if they are mutually orthogonal during the symbol time. This is valid if the frequency spacing between successive carriers is $\Delta f = 1/T_S$. In HD-DIVINE the symbol time T_S is 64 µs (Figure 8).

To make the modulation system even more robust against intersymbol interference, another method combined with the long symbol time is used. The symbol time T_S is stretched to cover a small additional period ΔT where no information is sent. The period ΔT is known as the "guard interval" and, if given a proper value, this totally eliminates the intersymbol interference. In HD-DIVINE a guard interval with $\Delta T = 2$ µs is optional (Figure 9).

The use of a guard interval will avoid inter-symbol interference but the phase and magnitude of the received signal will still be affected. This problem is solved by dynamic equalization. A pre-determined signal with a known phase and amplitude is sent regularly, and the receiver uses this signal to measure the influence of the channel. An equalization characteristic for each carrier is calculated on the basis of this measurement. In this way the receiver can, theoretically, compensate for the channel perfectly. The method used required that the channel is stationary, i.e. is stable over a period of time, which should be the normal case for HDTV reception. As the channel is impaired by noise an averaging procedure is done in order to get a more accurate measurement; this also demands a stationary channel.

Each single carrier in the prototype system is modulated using 16-state quadrature amplitude modulation (16-QAM), i.e. each symbol transmits four bits. As the total length of a symbol T is $T_S + \Delta T$ = 66 µs and the total number of carriers used is 448, the usable bitrate is 448 × 1 / $T_S + \Delta T$ = Mbit/s. The error-correction code is a shortened Reed-Solomon code (208,224) which leaves a net bitrate of 27 × 208 / 224 = 25 Mbit/s. The modem is, strictly speaking, not COFDM but OFDM, in the sense that the forward-error correction is not incorporated in the modem.

However, the overall system implementing COFDM as the forward-error protection is incorporated in the video source coder/decoder.

The modulated carriers are generated using a 1024-point inverse discrete Fourier transform (IDFT) implemented using a fast Fourier transform. With the use of modern integrated circuit technology the modem has a very small physical size (the modulator has just three boards, each measuring 25 x 63 cm).

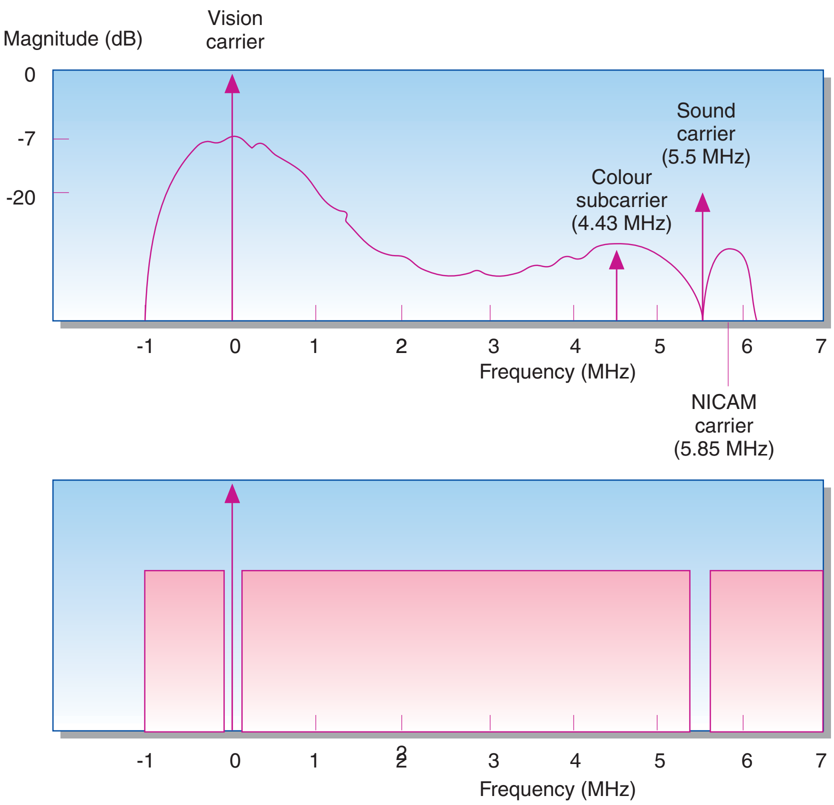


Figure 7 Spectrum of PAL and OFDM signals

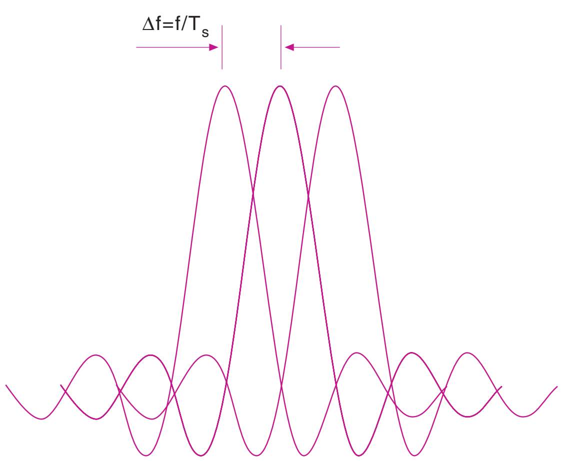


Figure 8 Overlapping spectra of three closely-spaced carriers in an OFDM system

2.6 Network planning

The network planning studies conducted in the HD-DIVINE project are based on the assumption that directional receiving antennas will be used. This simplifies frequency planning considerably for conventional as well as single-frequency networks. This makes it possible to use the same channel for two different programmes, with a minimum of separation between the two. This is a vital difference compared to the situation in DAB, where non-directional antennas are used. The protection ratio for OFDM versus OFDM (16-QAM) is virtually the same as that which can be expected from directional receiving installations. The guard interval completely eliminates the effect of intersymbol interference when the time difference between the OFDM signal and its echoes is less than the length of the guard interval. The negative effect of selective fading can be eliminated by proper channel coding.

Good reception may require a carrier-to-noise ratio (C/N) of less than 20 dB; this is an improvement of more than 20 dB compared to analogue PAL. Hence, the transmitted power for a given coverage area can be reduced significantly. With techniques for integrated coding and modulation (e.g. trellis-coded modulation) the required C/N ratio will be even lower. Planning is greatly simplified because the protection ratio between the wanted OFDM signal and an unwanted OFDM signal is about an order of magnitude (i.e. 20 dB) better than the best protection ratios obtained using precision offset in PAL channels. When the time difference between the direct and delayed paths is within the guard interval, the orders of magnitude of the signals on each path may be equal, without any risk of inter-symbol interference. As the time difference increases the magnitude of the delayed signal decreases and thus reduces the effect of inter-symbol interference.

One attractive way of using the single-frequency option is in the form of single-frequency coverage extenders, or gap fillers, in connection with in-band networks using conventional frequency planning. In the same way, a regional power profile can be designed to fit the desired coverage area. The basic idea is to increase the number of transmitter sites and to reduce the total emitted power. The lack of overspill reduces the frequency reuse distance to a minimum. This power profiling technique will facilitate the introduction of single-frequency services, with a minimum of blocks available for single-frequency networks throughout Europe. Thus the flexibility in planning networks for single-frequency operation opens up the possibility of a combination of national, regional and local programmes, without adversely affecting frequency economy.

During a transition period, analogue television and digital HDTV must use the same part of spectrum, possibly in the form of simulcasting. In this period the coverage of digital HDTV will be limited more by interference from analogue television transmitters than from digital television or noise. An important measure to reduce the effect of interference from analogue television is to leave unused the OFDM carriers that are subject to interference from the vision and sound carriers of the analogue system, as described above.

Planning exercises suggest that the simultaneous broadcasting of digital HDTV and current analogue services is possible, even in areas with intense use of the spectrum, such as the Öresund region between Sweden and Denmark. This is made possible by the low power used to transmit digital HDTV, the use of directional antennas and the use of different polarizations.

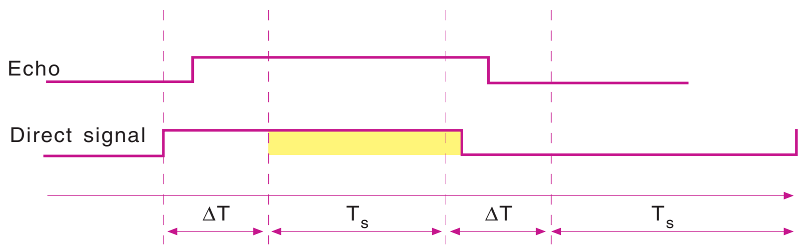


Figure 9 Use of the guard interval, ΔT, to eliminate intersymbol interference

3 Conclusions

HD-DIVINE aims at attaining a balanced solution for broadcasters, audiences and the electronics industry for the next generation of television broadcasting standard. It is our hope that the main principles of HD-DIVINE will guide the process of international standardization. The system as it stands today is not complete. It is the first working prototype. During the next two years, efforts within the project will focus on a final system. It is our belief that, to be a success, a new television broadcasting standard must offer something to all parties concerned.

References

1 Appelquist, P. The HD-DIVINE project: A Scandinavian Terrestrial HDTV System. *NAB HDTV World Conference Proceedings,* April 1993.

2 Appelquist, P, Franceschi, O. Development of a Terrestrial Digital HDTV System: The HD-DIVINE Project. *EBU Technical Review,* 253, 1992.

3 *Report to the European Launching Group on the Prospects for Digital Terrestrial Television,* WGDTB 1063. November 1992.

4 Stare, E. Development of a Prototype System for Digital Terrestrial HDTV. *TELE,* English edition, 1992.

5 Mason, A G, Lodge, N K. Digital Terrestrial Television Development in the SPECTRE Project. *IBC Technical Papers,* Conference Publication, 358, 1992.

6 Bernard, P. STERNE: the C.C.E.T.T. Proposal for Digital Television Broadcasting. *IBC Technical Papers,* Conference Publication, 358, 1992.

7 Mason, A G, Drury, G M, Lodge, N K. Digital Television to the Home: When will it come? *IBC Technical Papers,* Conference Publication, 327, 1990.

Optical Routing Systems in Radio and Television Production Centres

BY BJØRN M ELGSTØEN

1 Introduction

Most Radio and TV Production Centres have one or more switching matrices for the routing of audio and video signals between the production facilities and from there to the transmission network. These matrices are built for analogue signals.

A new era has now started with the introduction of digital and PC based production equipment. Outside the Production Centres the transmission networks are being digitized. The switching matrices are therefore no longer optimal for their purpose and, even worse, they are quality reducing because of all the analogue/digital/ analogue transformations they incur.

It is evident that new routing systems will be designed for digital signals. Replacement of the existing matrices will also involve replacement of the cabling system, as much of the existing cables are unsuited for digital signals. New optical fibres will be required. The task we are faced with is therefore a completely new digital infrastructure.

Figure 1 shows, very much simplified, the main functions and the main signal routes in the NRK Radio Production Centre in Oslo. Most public broadcasters have a structure similar to this. The switching matrices may be less comprehensive in some places and the manual routing functions in the switching centre more predominant. In Oslo, there are at present 3 continuity rooms which can provide for the program feed to the transmitters.

Figure 2 shows, similarly simplified, the main audio and video routes in the NRK TV Production Centre in Oslo. As a result of the price, the size and the complexity of the videotape and film machines at the time when the Production Centre was built, the machines were assembled in a central area as a common resource. The complexity of this area made it necessary to localize big switching matrices here. Otherwise there are small switching systems both in the switching centre itself and in the continuity suite. Most of the routing to and from external lines are made manually in the switching centre.

Decisive for the development of Radio and TV production techniques are:

- The digitizing of the audio and video signals, and the effective compression techniques being developed
- The enormous progress made in data processing and storage.

The difference between commercial multimedia equipment and professional audio and video equipment diminishes. The Broadcasters should welcome this trend, which will lead to cheaper equipment.

New computer based production facilities for Radio and TV will replace analogue facilities and offer a lot of new services. They will necessitate production routines which will hopefully make a more rational use of the production staff. We are already seeing that the structures shown in Figures 1 and 2 are being reorganized. This tells us that a new infrastructure can end up looking quite different from the existing one, and it becomes important to choose solutions as flexible and open as possible.

2 Routing in a Radio Production Centre

Analogue infrastructures have inherent problems like electromagnetically induced hum, noise and crosstalk. With the introduction of digital infrastructures these problems will disappear almost completely. However, a new category of problems arises and must be thoroughly discussed. Points to be discussed are:

- Digital audio formats
- Synchronization
- Delay
- Transmission in real time
- The role of the technical switching centre; emergency.

Digital audio formats

After many years with non-compatible and non-standard audio formats it is now possible to choose standardized formats for the in-house use.

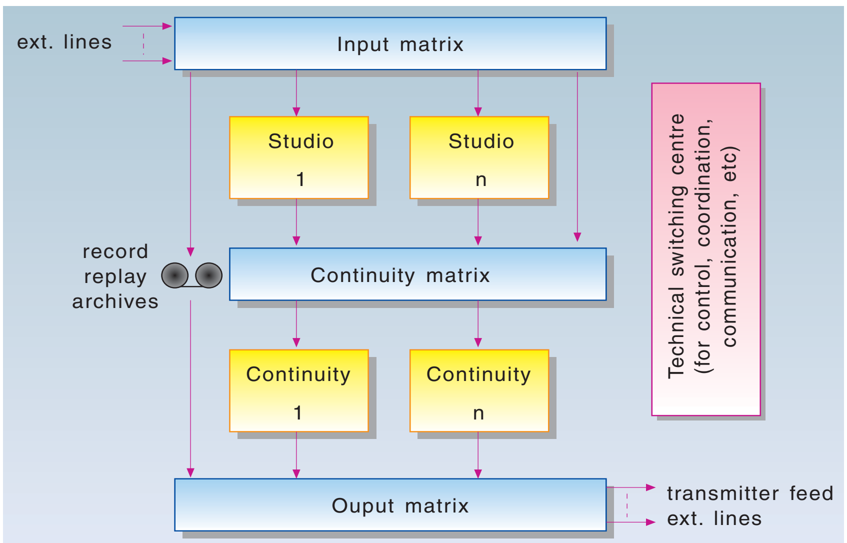


Figure 1 Typical analogue infrastructure in a Radio Production Centre

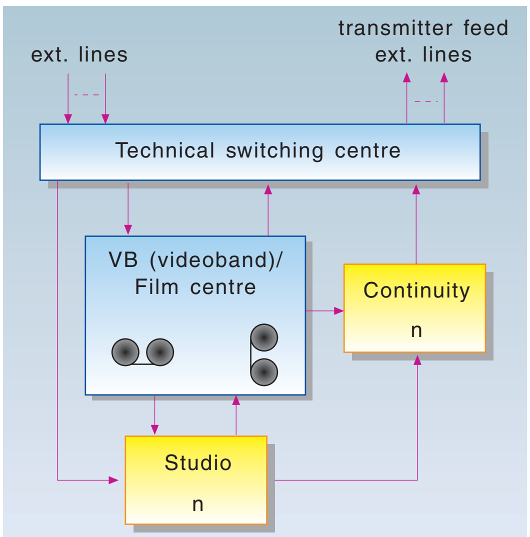


Figure 2 Typical analogue infrastructure in a TV Production Centre

The CCIR studio standard must be the basis which specifies a sampling frequency of 48 kHz and a word length in the range of 16-24 bits. The AES/EBU series interface for stereo (2 channel) transmission in real time will be used. Transmission data rate is 3,072 Mbps. (EBU Rec. 3250, AES 3-1992, CCIR Rec. 647, IEC 958.)

Multichannel transmission in real time shall conform to the recommendation AES 10-1991(ANSI S4.43-1991), usually called MADI (Multichannel Audio Digital Interface). The MADI format is a time division multiplex of 56 mono sound channels. There is compatibility between MADI and the AES/EBU format, so that 28 AES/EBU signals fit exactly into one MADI. The MADI transmission data rate is 125 Mbps.

Bit reduced audio shall follow the ISO/MPEG 11172, layer II recommendation (Musicam). The data rate for a stereo channel can be selected from 384 to 128 kbps. 384 kbps will offer excellent quality. A bit reduction down to 128 kbps is possible with only minor perceptible quality impairments. ISO/ MPEG layer II has been selected to be used in digital audio broadcasting (DAB). Compression can be performed directly from the studio standard and vice versa. ISO/MPEG, layer III is an interesting recommendation for very low data rates, e.g. on ISDN lines. This standard will be transformed to the layer II standard or the studio standard.

Synchronization

Audio signal mixing requires signals which are synchronous. A reference clock signal must be available at the mixing point. Specifications for such a clock are given in the AES Rec.11-1991. External audio signals must be resynchronized to this clock.

There must be a centrally situated master clock in a multistudio environment. Every studio must have a local clock, slaved to the master clock. This will achieve synchronization when two or more studios shall interact. At the same time the studios may function separately if the master clock fails. This configuration may be modified if a central, synchronous router is installed. The synchronization between studios is in that case maintained via the router.

The Radio master clock will probably be synchronized to a corresponding master clock in the TV Production Centre. This enables co-production between Radio and TV. In due time, when SDH Tele-networks are in operation, it may be appropriate to synchronize the Production Centres to the Tele-network.

Delay

Digital processing takes time. A/D transformation, bit compression, expansion, multiplexing, etc. all incur some time delay compared to the original sound. Disturbing echo-effects may occur when listening to sound signals received via two different program chains. Reporters using headsets have already experienced the annoying situation when they hear their own delayed speech. Program discussion sounds strange when the conversation takes place via lines with a delay. In a TV environment different delays along the audio and the video signal path may end up in the well known situation of non synchronization between sound and picture. This must be considered when sound and picture are treated separately (post production).

Transmission in real time

A real time transmission lasts exactly as long as the program material itself. It is obvious that live transmissions must take place in real time. A consequence of this is that at least some future routing must be in real time. In an analogue situation all routing is real time routing. However, with digitized program material and no specific requirements for a real time transmission, the possibilities for a faster data transfer (to save time) or a slower data transfer (e.g. on ISDN lines) will be most welcome.

The role of the technical switching centre; emergency

At present the technical switching centre can control the quality of the essential connections through the switching matrices. In an emergency situation the switching centre staff can easily manually bypass the failing link.

In an analogue infrastructure, sound monitoring with quality control can be done almost everywhere, just by tapping the line. In a digital infrastructure this is not feasible, as the signal transport may be by data file or in a multiplexed format. Even tapping of a single line with an AES/EBU signal is difficult without affecting the signal on the line.

Manual patch field operation as a possibility in emergency situations will also become difficult. Connectors for single mode optical fibres are on the market, but they must be treated with special care, and for the time being they are not robust enough to be used in a patch field.

As a consequence, the influence of the technical switching centre will be somewhat reduced as far as internal routing is concerned. Outgoing and incoming PTT lines should continue to be controlled in the technical switching centre.

2.1 Routing based on switching or multiplexing

Switching matrices are available which are capable of switching both analogue and digital signals. They can be attractive in a period where both analogue and digital signals are present, but they are, in principle, analogue matrices and will not offer the best alternative for digital signals in the future. Optical fibre technology is interesting even for audio switching because of the immunity from electromagnetic hum and noise.

A system is shown in Figure 3 based on time division multiplexing, developed as a prototype by the BBC around 1990. It is based on the MADI recommendation. Four 125 Mbps MADIs are multiplexed, giving 500 Mbps. This signal is transmitted via an optical fibre to a central optical splitter. The 16 splitter outputs are then distributed back to the studios on optical fibres where each one contains all 224 monochannels (56 x 4).

However, for the time being it seems impossible to build an optical system for pure radio like the one described at an acceptable price. Presently available commercial MADI routers do not use expensive electric multiplexing/ demultiplexing at data rates above the MADI level. Instead all the MADIs are taken to a central switching system on optical fibres. Transceivers developed for data networks (FDDI) are used to feed the optical fibres.

A MADI-based routing system to handle real time transmissions probably will be a good choice. This will be valid even if we build data networks as described

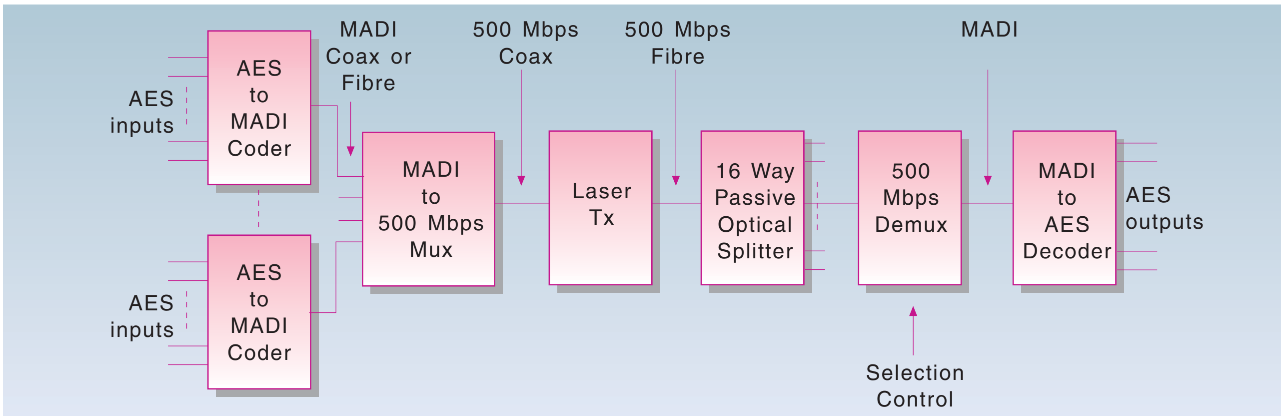


Figure 3 Multi-channel Audio Digital Interface (MADI) system

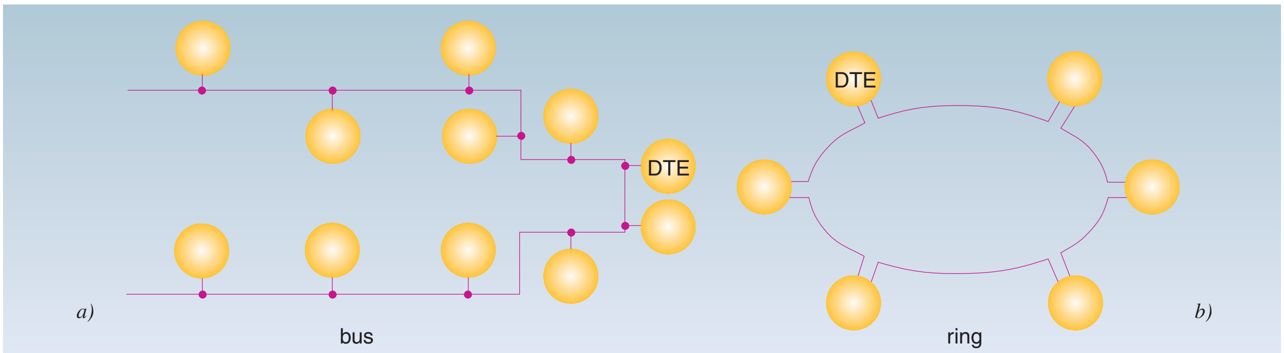


Figure 4 LAN topologies

later, simply because the data networks cannot present enough capacity to manage all the routing requirements in a Radio Production Centre like NRK. A MADI system can operate in a synchronous mode and switch signals "on air".

2.2 Routing based on data networks

With the introduction of PC based facilities for audio editing and storage, the idea of using a data network to interconnect the units naturally arises.

Most Radio and TV stations already have some sort of data network installed for administrative services. The network cabling does not necessarily include the technical areas and the studios, but widening the network to include these areas would normally be easy.

The two types of LANs in common use are:

- Ethernet
- Token Ring.

Ethernet according to IEEE recommendation 802.3 uses a Bus topology (Figure 4a.) The recommendation originally specified the use of coaxial cable, but today a combination of optical fibres and twisted pairs is more commonly used.

The access method in 802.3 is called CSMA/CD (Carrier Sense Multiple Access with Collision Detection). All stations connected to the bus listen to the cable to detect if a data frame is currently being transmitted. If this is not the case, one of the stations may start to send data. The data will be encapsulated in a frame with the destination address at the head of the frame. The frame size can vary between 72-1676 bytes. If a transmission – i.e. a carrier signal – is sensed, the sending station defers its transmission until the network is free. Even so, two stations wishing to transmit a frame may determine that the network has been free at the same instant of time and hence both start to transmit simultaneously. A collision is then said to occur and the contents of both frames will be corrupted. The collision will be detected by both sending stations. They will then wait for a random time interval before trying to retransmit the affected frames again.

The Ethernet bit rate is 10 Mbps, but due to the access method only 30–40 % of the capacity can be utilized.

Token Ring according to IEEE recommendation 802.5 uses a Ring topology (Figure 4b) The access method is called token passing because a token packet circulates around the ring when all stations are idle. A station with data packets to transmit must wait until it detects a passing token. The station then captures the token by aborting the token transmission. After it has completely received the captured token, the station begins transmitting one or more packets. The packets are picked up at the destination, and the transmitting workstation releases the token by inserting a new token on the ring. Other stations with data to send must wait until transmission is completed and the token is released. A typical Token Ring packet size is 132 bytes.

LANs are used for distributed control of video and audio switching matrices. For such purposes a LAN is well suited. Contrarily, a LAN is not well suited for the transmission of a real time audio signal.

A digital audio signal is quite different from the data frames or packets which are found in LANs. The bit rate for a real time audio signal is directly related to the sampling frequency, the sample word length, the number of channels etc. Digital audio needs some sort of “circuit emulation” through the network, i.e. a continuous, constant data stream is required. A transmission based on these requirements is often called isochronous.

Isochronous transmission does not fit into the LAN recommendations, where data packets are sent irregularly and the packet size is allowed to vary. Neither does the LAN data capacity suffice for audio transmission. This is quite obvious, bearing in mind that one AES/EBU channel requires 3 Mbps for a real time transmission.

LANs can be used, however, for file transfer between workstations when real time transmission is not necessary, or when a delay can be accepted. Some of the latest audio editing systems take advantage of a bit compressed audio format (i.e. Musicam). Provided there is a limited number of simultaneous transmissions, a LAN can be used, but heavy buffering will be necessary in the workstations, and bursty traffic due to other data services must be kept out of the network.

Audio file transfer between equipment from different manufacturers has up to now been a problem, as no standard exists.

OMFI (Open Media Framework Interchange) seems to define a framework that can overcome the lack of file transfer compatibility. The proposed OMFI has a broad support from the manufacturing industry, and thus has become a “de facto” standard.
OMFI aims at the whole multimedia range. Audio, video, graphics and pictures can be transferred together with data for combining and presenting the media.

2.3 FDDI (Fibre Distributed Data Interface)

FDDI is a high speed LAN (100 Mbps) specified and standardized by ANSI. The access method is token passing, like in Token Ring. In addition a network control is provided, enabling errors to be isolated.

FDDI has not yet been a big success, partly because of the price, and partly due to the fact that the increased capacity still is not needed by most user groups. For the purpose of audio routing FDDI represents no improvement compared to an ordinary LAN. Isochronous transmission is not possible.

To remedy this weakness, the ISO has defined FDDI II, which is intended to satisfy multimedia requirements.

FDDI II is an upward-compatible extension to FDDI that provides a circuit switched service while still maintaining FDDIs token controlled packet switched service. A frame structure is repeated every 125 s and circulated around the ring, i.e. with a frequency of 8000 Hz. An isochronous transmission will be given access to regularly repeating time slots in the frames.

Figure 5 shows the FDDI II architecture with the link layer and the physical layer of the OSI model. The most interesting are the two new components added to the token passing MAC module at the MAC level: The HMUX (Hybrid Multiplexer) and the IMAC (Isochronous MAC). The HMUX multiplexes packet data from the MAC and the isochronous data from the IMAC.

An FDDI II network can operate in either a “basic mode” or a “hybrid mode” In basic mode only the packet switched service is available (same fashion as FDDI). In the hybrid mode both packet switched and circuit switched services are available.

The framing structure of FDDI II is shown in Figure 6. The frame consists mainly of 16 WBCs (Wide Band Channels), each of them with 96 octets. That makes a data capacity of 6,144 Mbps per WBC (96 x 8 x 8000) This covers exactly two AES/EBU channels, each of them 3,072 Mbps. The 12 DPGs on the left side of the frame contain one octet each, and they are used only for token controlled data. Including preamble and header (116 bits) the whole frame contains 12500 bits (12500 x 8000 = 100 Mbps). Each channel (WBC) can be set aside for circuit switching or packet switching exclusively. A circuit switched channel can be subdivided into several individual, isochronous channels.

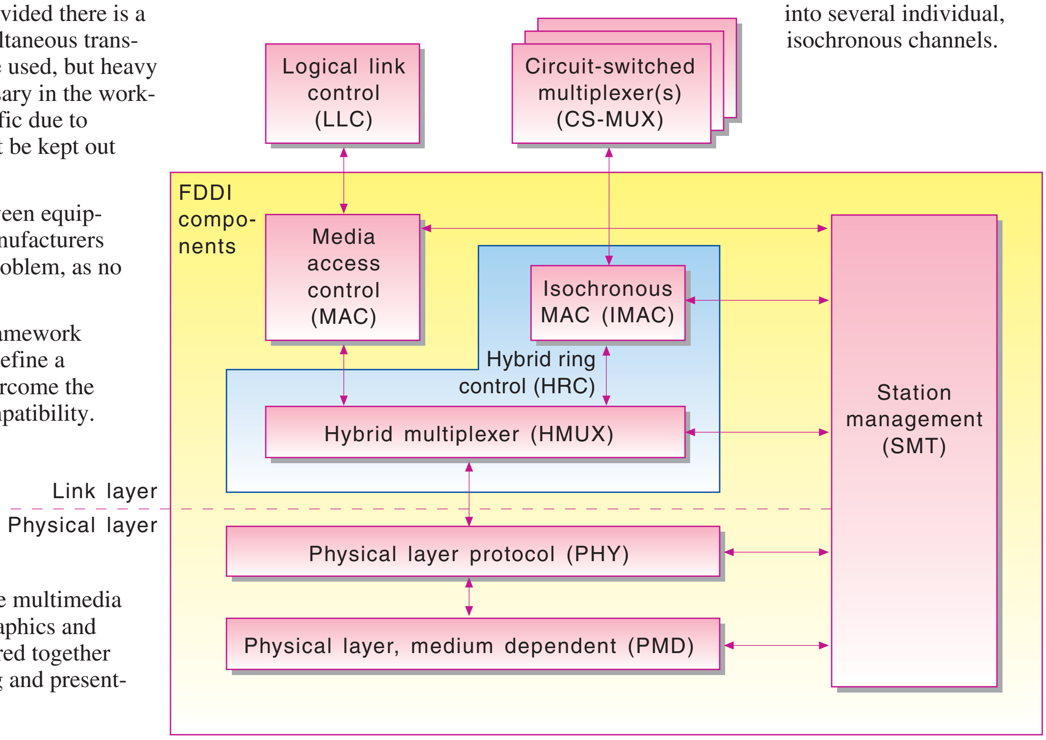


Figure 5 FDDI II architecture

In an FDDI II a station called a cycle master generates the framing structure 8000 times per second. The cycle master strips each cycle as it completes its circuit around the ring.

2.4 High speed LANs

In parallel with the development and standardization of FDDI and FDDI II some of the giants in the data industry have started work to propose a new Ethernet standard operating at a much higher speed. An FDDI installation can be quite expensive since an optical fibre ring is required . An important aim for the work on the new Ethernet standard has therefore been that existing cabling (i.e. coaxial cable and twisted pair of category 3 or better) still can be used.

The work has ended up in two proposals: 100 VG Any LAN and 100 Mbps CSMA/CD.

100 VG Any LAN is planned to replace both Ethernet and Token Ring in the future. The access method resembles token passing, but without the use of a token. Any LAN is thought of as an alternative even for multimedia users with a requirement for isochronous transmission. It remains to be seen if 100 VG Any LAN will be an alternative for professional audio routing.

100 Mbps CSMA/CD (Fast Ethernet) is simply an Ethernet with the speed increased from 10 – 100 Mbps. The access method (CSMA/CD) is unchanged, and we know that it is not well suited for real time audio transmissions. However, used for file transfer between workstations Fast Ethernet can be an interesting alternative.

3 Routing in a TV Production Centre

The revolution in digital video signal processing has now reached a level where digital "islands" are frequently found in the analogue surroundings. Picture manipulation and graphical work are now done by means of digital signal processing. Digital videotape machines and digital mixers are ordinary equipment on the market, and digital TV signals will soon be received via satellite.

There is already a need for transmission of digital signals between the "islands". Next step will be a digital routing system with a correspondingly reduced need for analogue routing. In NRK we assume that a transition to an all digital TV house will go on until the year 2000.

The digital TV signal requires a large transmission capacity (100 times more than a digital stereo sound channel). It is therefore likely that a TV Production Centre will have an optical infrastructure. We are now in an exceptional situation. Specifications for a new routing system can be drawn up which would not be possible to comply with before:

- Non-electric connection, i.e. optical connection, with immunity from hum and other types of electromagnetically created noise
- No need for signal correction within a distance of 1 – 2 km
- All sources accessible at all destinations simultaneously, a flexibility not necessarily being fully utilized
- Source selection shall be possible at the user position (the destination)
- The system must be expandable in different degrees, and provide capacity for the future expansions
- The routing system shall be transparent, i.e. independent of videotape formats or standards on the user side
- The system shall be non-blocking. A routing set-up shall not disturb connections already in use.

The digital component video signal according to CCIR Rec. 601 has become a widely accepted studio standard. Almost all types of professional video equipment have today

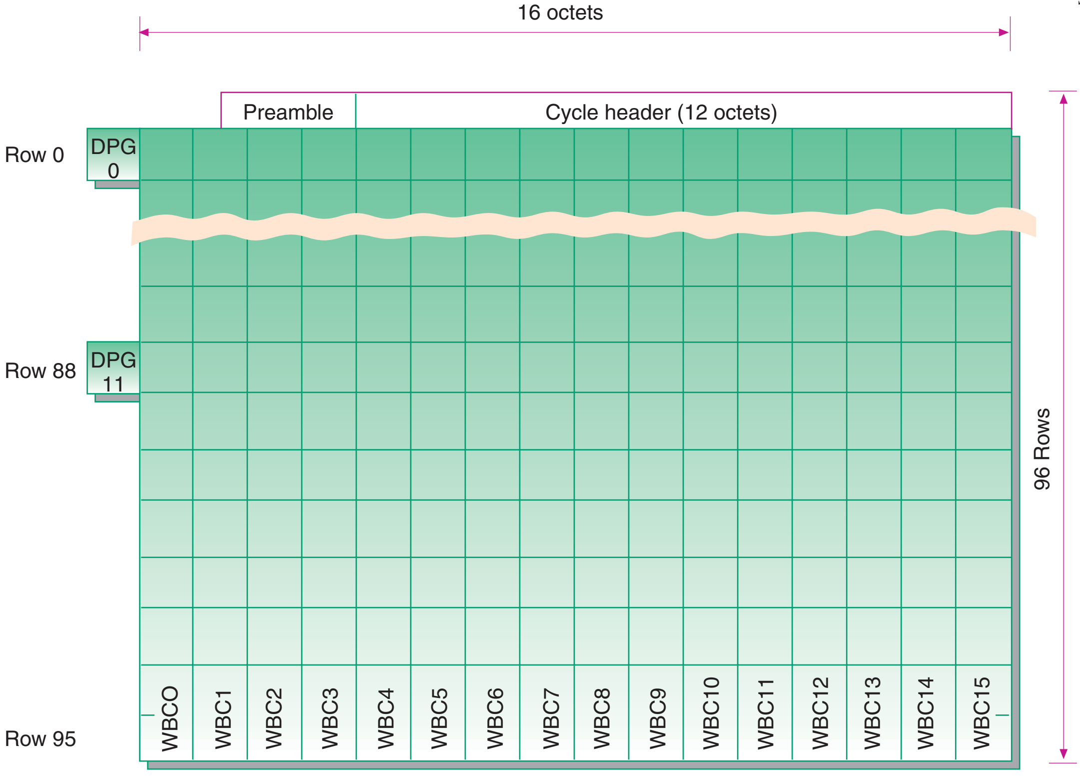


Figure 6 FDDI II framing structure

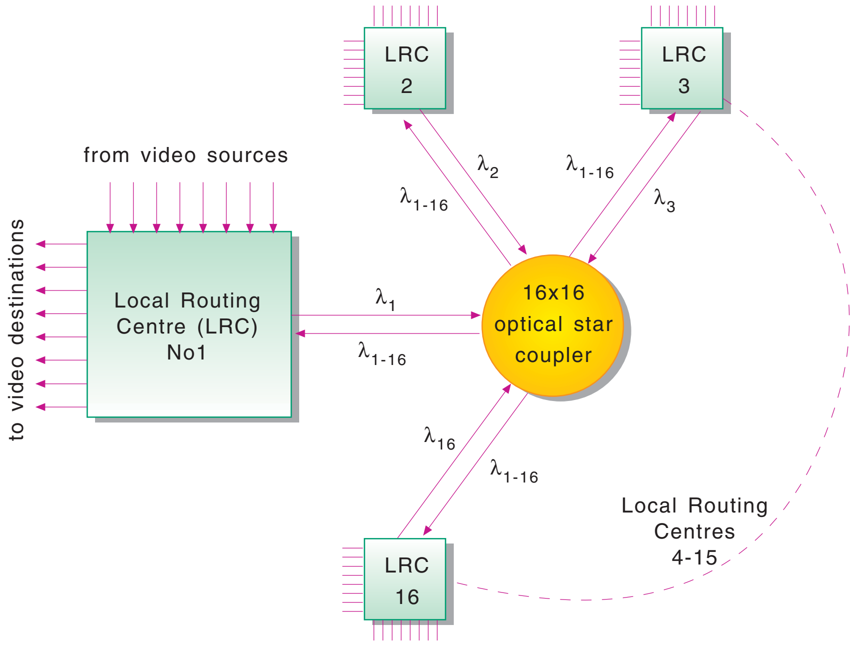


Figure 7 WTDM routing system

inputs/ outputs adapted to the 601-format. Therefore, the series format of the 601-signal (CCIR Rec. 656) will be a natural choice for the interface between sources and destinations and the routing system itself. The serial interface data rate is 270 Mbps.

Composite video (PAL) should, if possible, be avoided in the new routing system.

In addition to the video signal, the 270 Mbps stream has enough space for a high quality stereo sound channel (preferably the AES/EBU format), communication and signalling functions. However, NRK will try to standardize on a commonly used interface for video with embedded audio, and drop the communication and signalling functions due to the costs. Signalling functions are no longer necessary, and communication can take place in an ordinary intercom system outside the routing system which would be necessary anyway.

3.1 RACE 1036, WTDM Broadband Customers Premises Network

The aim for RACE (Research and development in Advanced Communications technologies in Europe) is to develop the IBC or IBCN (Integrated Broadband Communications Network). Within this broadband network some user groups have been focused on because of their supposed need for broadband capacity, and because they have special requirements for the network. The TV companies make up an important part of the user group of the IBC. One of the RACE projects (RACE 1036) started in 1988 to study possible ways to apply optical techniques to the routing needs in TV Production Centres. The studies were based on work done by the BBC Research Department and a system proposal was presented as early as 1985. The work in RACE 1036 has been led by the BBC.

The new system, verified and demonstrated in RACE 1036, is based on a network of local routing centres (LRCs), each serving a number of sources and destinations (see Figure 7). In each LRC an SDH time division multiplexer combines 16 STM-1 levels (155 Mbps each) to form an STM-16 level of 2.5 Gbps. This 2.5 Gbps signal is fed to a semiconductor laser which turns it into an optical signal at a closely defined wavelength. The wavelengths are spaced by 4 nm in the 1500 nm band. Since a component video signal (Rec. 601) needs 270 Mbps, two STM-1's must be combined to transport one video program. That means 8 video signals in 2.5 Gbps (see Figure 8). At the centre of the system the optical signals from all the LRCs are

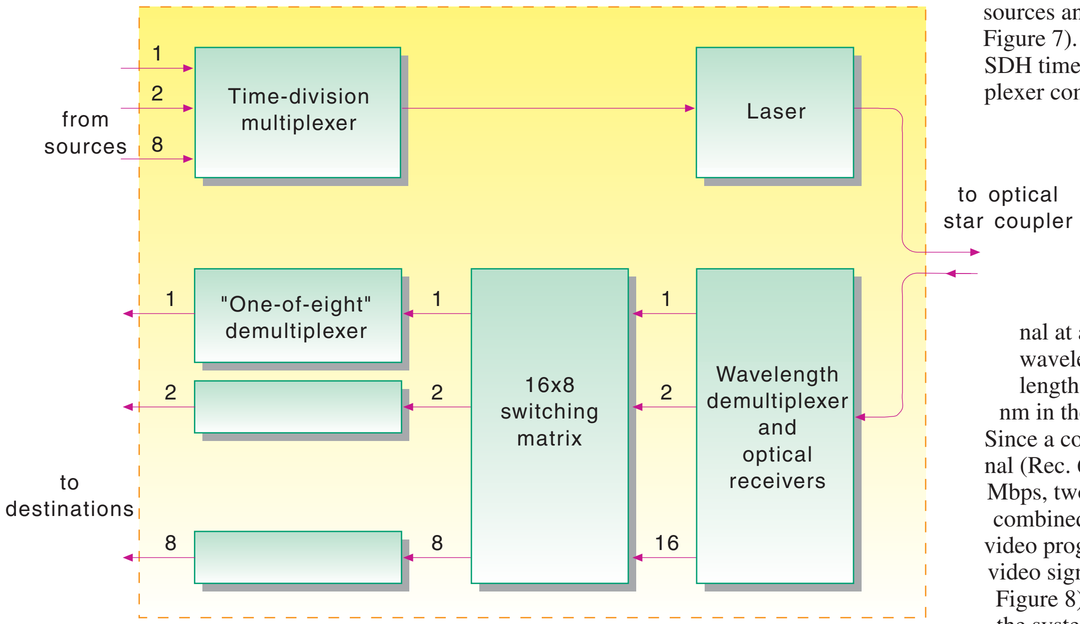


Figure 8 Local Routing Centre

combined in an optical star coupler which distributes the combined (optical wavelength multiplexed) signal back to all the LRCs. Each destination can select any source by first selecting the right wavelength and then selecting the desired signal from those available on that wavelength (see Figure 8).

This system has been known as a WTDM system because multiplexing both in the optical and the electrical domain has been utilized (WTDM = Wavelength and Time Division Multiplex). WDM should be preferred as the name of the optical multiplexing function instead of OFDM (Optical Frequency Division Multiplexing), which is used in some literature on the subject. OFDM is too close to the abbreviation used for the modulation method in digital broadcasting, COFDM.

The WTDM system is originally designed as a passive system, due to the fact that optical amplifiers and optical jackfields have been difficult to realize up to now. To avoid the use of optical amplifiers, it has been necessary to look very closely at the losses in every link of the system, and so reduce the losses as much as possible.

The lasers must deliver a high output power (typical -4 dBm) and the optical receivers must have a high sensitivity (typical -29 dBm). The total losses through the optical chain must be kept below this difference (25 dB).

In November 1992 the Project demonstrated its final test bed. The demonstration showed that all the aspects of a WTDM system are technically feasible. The system was also successfully demonstrated at the TV Symposium in Montreux in June 1993.

The star configuration shown in Figure 7 has one significant weakness: If the star coupler breaks down, the whole system is dead. To overcome this, alternative system architectures have been proposed. One of them is shown in Figure 9.

This global architecture is based on 8 x 8 star couplers and a network of optical fibres. The architecture looks complicated, but both optical fibre and the star couplers are rather cheap elements. By using only 8 wavelengths another advantage is achieved: The system can be kept within the bandwidth for optical amplifiers (typical 1530 – 1560 nm), thus enabling new methods to be introduced in the optical demultiplexing process.

3.2 RACE 2001, the Pilot Installation

A new RACE project, R 2001, started work early 1992. R 2001 is taking the WTDM system demonstrated in R 1036 towards a pilot installation. NRK joined the R 2001 project as a sponsoring partner. Recently, this has been changed into a full partnership.

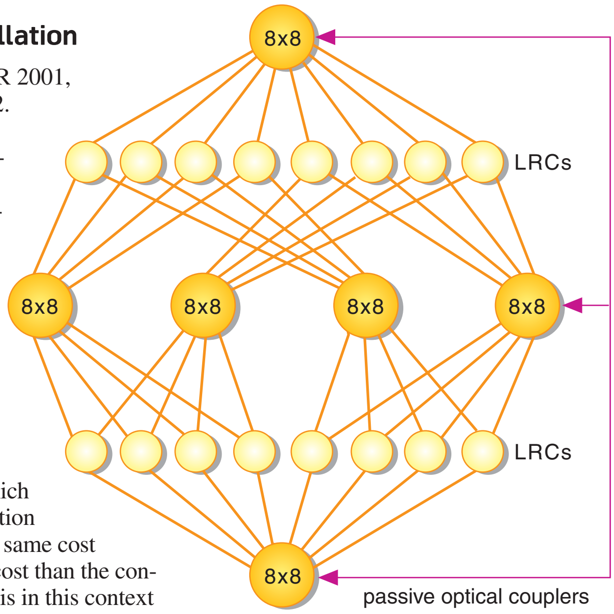


Figure 9 Global architecture with six 8 x 8 stars

Current work

It has become clear that any new technology which performs a routing function must be available at the same cost or, preferably, a lower cost than the conventional alternative. It is in this context that the marketable price of the WTDM router has been investigated. The costs of the various components which make up the proposed system are being analysed to ascertain what those costs are, and what opportunities exist to reduce them.

In this project DFB (Distributed Feedback) lasers have been used. In a laser transmitter it is the active device itself which is predominant in the cost. Major cost reductions can only be achieved by producing the devices in a more cost effective way. This is only viable when large manufacturing quantities are being considered, i.e. wider use of the optical multiplexing techniques.

The cost of passive optical components will be dramatically reduced by the use of silica-on-silicon technology, which is being developed within another RACE project, R 1008. This technology offers the prospect of mass production. A key feature is that v-shaped grooves for holding the input and output fibres can be precisely defined in the silicon substrate as part of the waveguide fabrication process. The silica-based waveguides offer both very low propagation losses and low fibre-to-waveguide interface losses. The fibre-to-waveguide alignment can be done passively and thus simple and cost reducing.

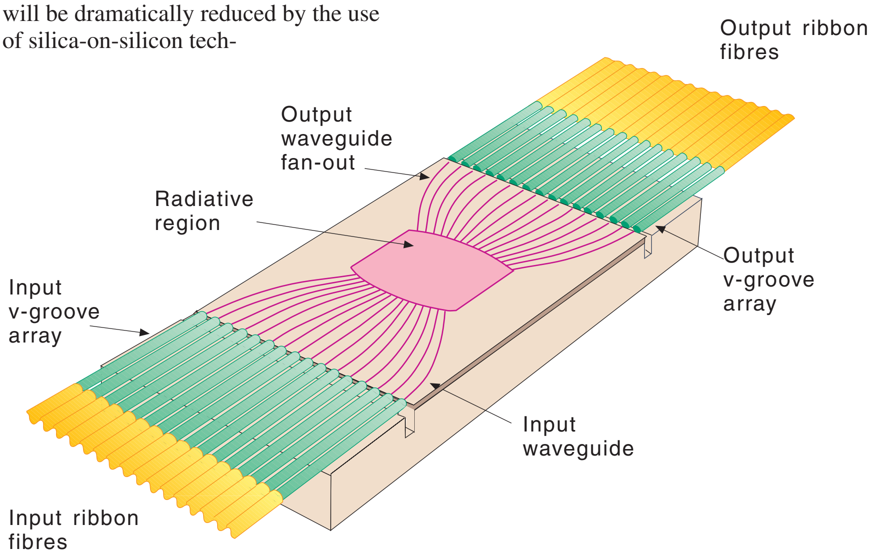


Figure 10 Radiative star 16 x 16 coupler (by permission from BNR Europe Limited)

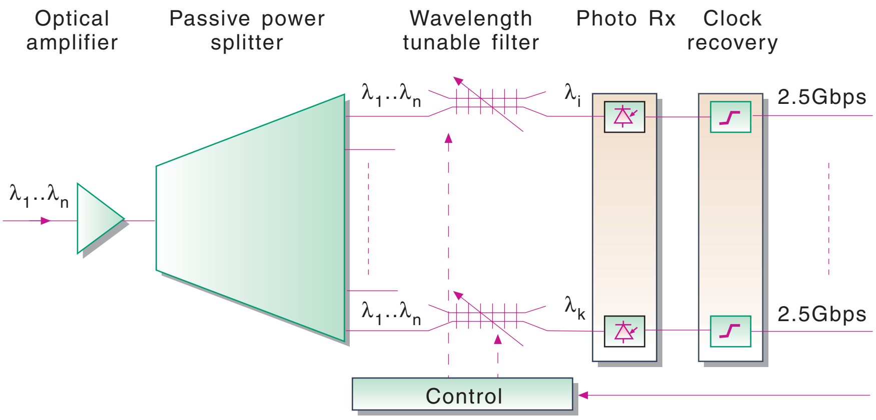


Figure 11 Optical switching based on EOTF filters (by permission from GEC Marconi)

A 16 x 16 star coupler based on silica-on-silicon technology is shown in Figure 10. Light from the input fibres is coupled via channel waveguides into a planar waveguide region, where it is allowed to radiate freely for a set distance. The light is then coupled simultaneously into 16 output waveguides which are arranged around the far field of the waveguide region. The radiative star coupler is wavelength independent within wide limits. Worst case port-to-port losses of 17 dB for the 16 x 16 star has been achieved.

A study of the equipment costs for the LRC (Local Routing Centre) has shown that the use of alternative technologies to achieve the same functionality as conventional technologies could give significant cost reductions. This is especially valid for the receiving end (see Figure 8). Important work has been done to replace the expensive silicon-based 16 x 8 electronic switching matrix with an integrated optical switching and wavelength demultiplexing device. This extended optical part would also reduce the number of optical receivers from 16 to 8.

In the first R 1036 demonstration optical demultiplexing was achieved by means of a diffraction grating and a focusing lens. Tuneable filters are now used to select a wavelength.

Two alternative optical switching architectures have been proposed. One is based on the design of a polarization-independent electro-optic tuneable filter (EOTF) on lithium niobate. The complete solution consists of a passive silica-on-silicon splitter, preceded by an optical (erbium fibre) amplifier, and followed by an EOTF filter (Figure 11). After the splitting of the wavelength multiplexed optical signal in a number of parallel ways towards the output of the system, each of the optical filters can select among all the wavelengths and thus perform the switching function. Detailed designs of EOTFs for 8 and 16 channels are now available (wavelengths at 4 nm or 2 nm spacing, respectively).

They have electronic control circuits developed to interface the optical switch to the network control system.

The second optical switching architecture which has been proposed will be a monolithically integrated device based on indium phosphide (Figure 12). This particular design is tailored for the system architecture shown in Figure 9, with six 8 x 8 stars. In this case each LRC receives wavelength multiplexed signals from three different stars and each destination has to select one of them. The active combiner is in fact a space switch.

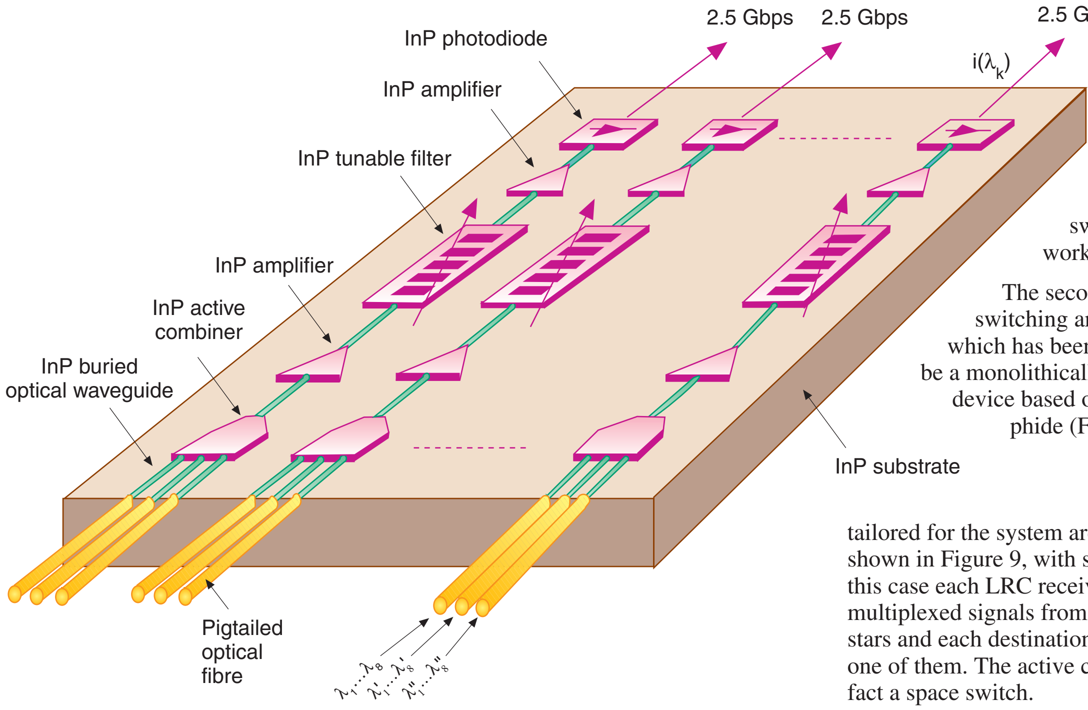


Figure 12 A fully integrated switch and optical receiver on indium phosphide (by permission from Thomson-CSF, France)

As some of the possible configurations on the demultiplexing side will need space switching, elements for this purpose are also being developed within the project. The elements are electro-optical switches produced on indium phosphide (InP), and two types are being developed:

- Directional couplers
- AOG (Amplifying Optical Gates).

The AOGs have the great advantage that they can be realized practically without losses. Switching matrices can be made by combining these elements on a hybrid form.

The silica-on-silicon technology offers many interesting possibilities in the design of passive components. A natural extension of the radiative star coupler work is the arrayed waveguide wavelength demultiplexer (Figure 13), basically two radiative stars joined by an array of waveguides of different lengths. Silica-on-silicon demultiplexers are close to realization, and in large scale production they would offer considerable cost benefits over existing demultiplexers. The configuration in Figure 13, a 1 x 16 demultiplexer with 4 nm channel separation would fit perfectly into the R 1036 architecture shown in Figure 7 and Figure 8.

Coherent technology is being developed in another RACE project. The intention is that R 2001 would provide an application in which this technology could be demonstrated. The coherent technique with optical heterodyne detection allows the optical wavelength spacing to be much closer (fractions of a nm), in fact only limited by the spacing that is required by the modulated signal. The capacity of the routing system can be considerably increased with this technique, which is especially interesting when HDTV is concerned. However, stabilization of the lasers can be difficult and costly, because the wavelength variations must be kept within extremely small tolerances.

Pilot installation

As the aim for R 2001 is to develop and install a commercially viable optical router based on the R 1036 proposal, a site for a pilot installation had to be

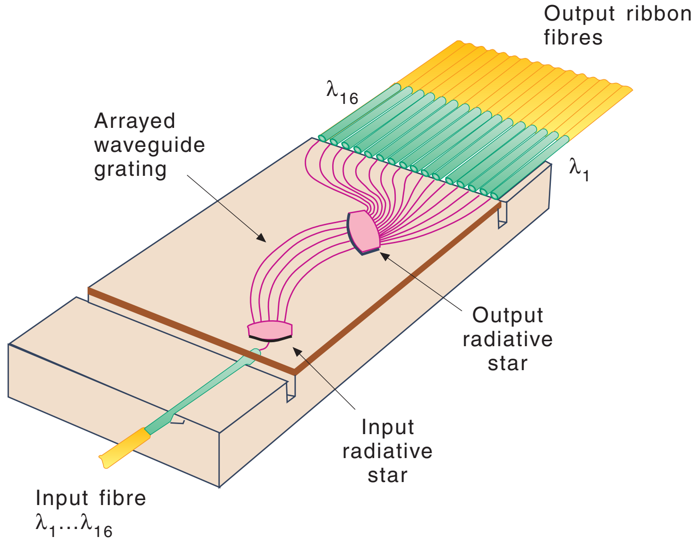


Figure 13 Arrayed waveguide wavelength multiplexer/demultiplexer (by permission from BNR Europe Limited)

Figure 14 OIU (Optical Interface Unit)
(picture taken at a demonstration in Brussels in 1993, reproduced by permission from GEC Marconi)

found. It has been decided that the site will be the NRK TV Production Centre in Oslo. One of the project members, GEC Marconi, England, will be the prime contractor for the entire system installation. The installation program will have four phases, which are designed to complement the existing modernization plans for the TV Centre. The start of phase 1 has been co-ordinated to start when the digital equipment used during the Winter Olympic Games at Lillehammer has been moved to the Production Centre in Oslo. The 4th phase is planned to complete the installation covering the whole Production Centre by the year 2000, but this will fall outside the project time scale.

The phase 1 installation will include the news studio (Studio 5), the graphical workshop area and the tape and film area. Each of these three nodes will have an OIU (Optical Interface Unit) installed. The OIU is shown in Figure 14. The OIU can house all actual optical and electrical modules needed in an LRC, and it has a unique construction with multipurpose electrical and optical backplanes. The unit has an interface to a network control system, which is also specified within the project.

In the pilot installation all the components and techniques described earlier in this article will be demonstrated and evaluated. With the OIUs it will be quite simple to change the configuration or expand the system in the future.

There is no doubt that a WTDM system like the one described would fulfil all relevant requirements for a TV routing system. The vital question will be the price, which is dependent on a large scale production of some optical components. Hopefully, the introduction of optical technology into CATV or related fields could create a market big enough for the component production.

4 ATM (Asynchronous Transfer Mode)

Experts claim that FDDI II most likely represents the end of the shared access technology. Next to come will be ATM, or cell relay. Quite soon ATM will be in general use in our data networks.

We know that ATM will be an essential part of the new telecommunication networks, as CCITT has chosen ATM for the switching functions and the transmission in B-ISDN. Many countries have already built pilot networks to gain experience with SDH, and even with ATM. The broadband routing system designed in R 1036 and R 2001 will have an interface against the IBCN outside the TV Production Centre. SDH will be common for both sides of the interface. When audio and video signals from the IBCN are received at the interface as ATM cells, it will probably be arranged to let the cells pass directly into the routing system.

Transmission of audio and video have always been a matter of circuit switching, with well known possibilities to protect the circuit and provide quality. Packet switching is something new for audio and video transport. ATM can take care of audio and video signals by "circuit emulation", i.e. giving the audio and video transmissions access to the cells regularly and thus ensuring a constant data rate with small delay. However, there is a delay in the network itself. Buffers will be needed both at the transmitting end and the receiving end. At the receiving end it will be necessary to accommodate jitter and resynchronize the signal. Error correction will most likely be added, as this is not provided in ATM. All together, ATM will cause some delay, but probably within acceptable limits.

There is some scepticism among broadcasting engineers to use a public SDH/ATM network for the transmission of Radio and TV programmes. They fear problems with faulty routing or overflow due to non-important and "bursty" services. The essential technical aspects to investigate will therefore be whether radio and TV transmissions can be separately controlled, and secured the necessary reliability.

Olympic Host Broadcaster Operations

BY ARILD HELLGREN AND HELENE AMLIE

The broadcasting of the Lillehammer Olympic Winter Games showcased the successful comprehensive use of digital technology. For the first time in Olympic history, the host broadcaster operation was based on an all-digital distribution and recording system. Still, most of the consumers, the world broadcasters, were operating in analogue. Compatibility with prevailing systems was thus an essential feature of the strategy adopted by the Olympic host broadcaster, NRK ORTO 94.

The Host Broadcaster Assignment

NRK's Olympic Radio and Television Organization, ORTO 94, was responsible for producing and directing the international radio and television coverage of the 1994 Olympic Winter Games in Lillehammer. The task was entrusted to NRK by the Lillehammer Olympic Organizing Committee (LOOC) early in 1990.

NRK ORTO 94 was established for the exclusive purpose of carrying out the most extensive project ever to be undertaken by a Nordic broadcasting company. In addition to producing the Olympic world feed – close to 350 hours of live radio and TV programme, distributed world-wide to an audience of more than two billion people – NRK ORTO 94 was in charge of planning, supplying and operating the technical and production facilities at the venues and the International Broadcasting Centre (IBC). A third major responsibility lay in the provision of requisite services and facilities to the rights-holding world broadcasters.

As an institution, NRK took on the task of host broadcaster for a fixed sum of 470 million Norwegian kroner paid by LOOC, to be applied solely to the financing of the host broadcaster activity. It was thus stipulated that ORTO was to be kept separate of the public broadcaster in both financial and administrative matters. The organization was accordingly formed with its own staff, budget and authority.

Project Stages

In early 1990, NRK ORTO 94 comprised 11 full-time staff members. That number grew to 81 by the end of 1993, and during the Olympics, host broadcaster operational staff totalled more than 1,300 people.

Major tasks during the first years of the project included planning of IBC and venue requirements. ORTO was involved from the very beginning in the design and construction of the Olympic sports venues, which were all, with one exception, built for the Games. ORTO's persistence in influencing construction to accommodate broadcasting requirements turned out to be vital to the success of the Lillehammer Olympics.

Throughout 1993, venue and IBC installations were completed and a series of sports events held at the Olympic venues to test the facilities, equipment and organization. In September 1993, the ORTO staff relocated from its Oslo headquarters to the International Broadcasting Centre to prepare the final pre-Games stage.

Having returned to Oslo, the host broadcaster has now begun debriefing and staff reduction. Final phase-out will be in July 1994.

Host Broadcaster System Design

The timing was right in 1990 for NRK ORTO 94 to give serious thought to digital broadcasting. Technology had advanced well beyond expectations and by 1991, the prospect of standardizing on serial digital interfacing began to materialize. After a thorough evaluation process, the decision regarding the overall system design for the Olympics was made in July 1992.

ORTO decided to apply the serial digital interface of CCIR Recommendation 656-1, with stereo audio embedded. The decision was influenced by the fact that Norwegian Telecom was constructing a digital optic fibre system for video and audio transmission from the Olympic venues to the IBC.

Serial digital routing ensured greater reliability and easier handling, and a stable signal quality was maintained through the entire transmission line. The installation volume was reduced substantially since the video signal with two audio channels was carried on one cable.

Combined with the choice of component Digital Betacam as the host broadcaster videotape format for recording and editing, the technical design allowed ORTO to stay in a digital domain from routing and transmission from the venues to archiving and editing at the IBC, without any loss of quality. An important consideration was Digital Betacam's compatibility with Betacam SP, the format used by the majority of the broadcasters' unilateral ENG crews.

After careful analysis, on-site product demonstrations and finally invitations for tenders, ORTO contracted Sony Broadcast International to supply the broadcasting system at the venues and the IBC. The contract stipulated the design, installation and maintenance of all stationary equipment to be operated by NRK ORTO 94.

The ORTO Olympic operation was on the whole based on hire. Some facilities, however, such as edit suites, VTRs and character generators, were purchased in cooperation with NRK Television and used by ORTO in the operational period. ORTO thus seized the opportunity to concentrate on solutions that would be of long-term benefit to NRK, as there was a wish within NRK to convert to digital component technology.

The installations supplied by Sony at the venues and the IBC were rented, like the mobile equipment brought by the production teams, microwave equipment, special cameras and so on. The estimated value of all stationary and mobile equipment deployed in the host broadcaster operation totalled approximately one billion Norwegian kroner (140 million USD).

The ORTO Production Teams

Given the structural makeup and size of the host broadcaster, agreements were reached with several broadcasting organizations to provide manpower and equipment during the Games. The European Broadcasting Union (EBU) and Canadian Television (CTV) contractually

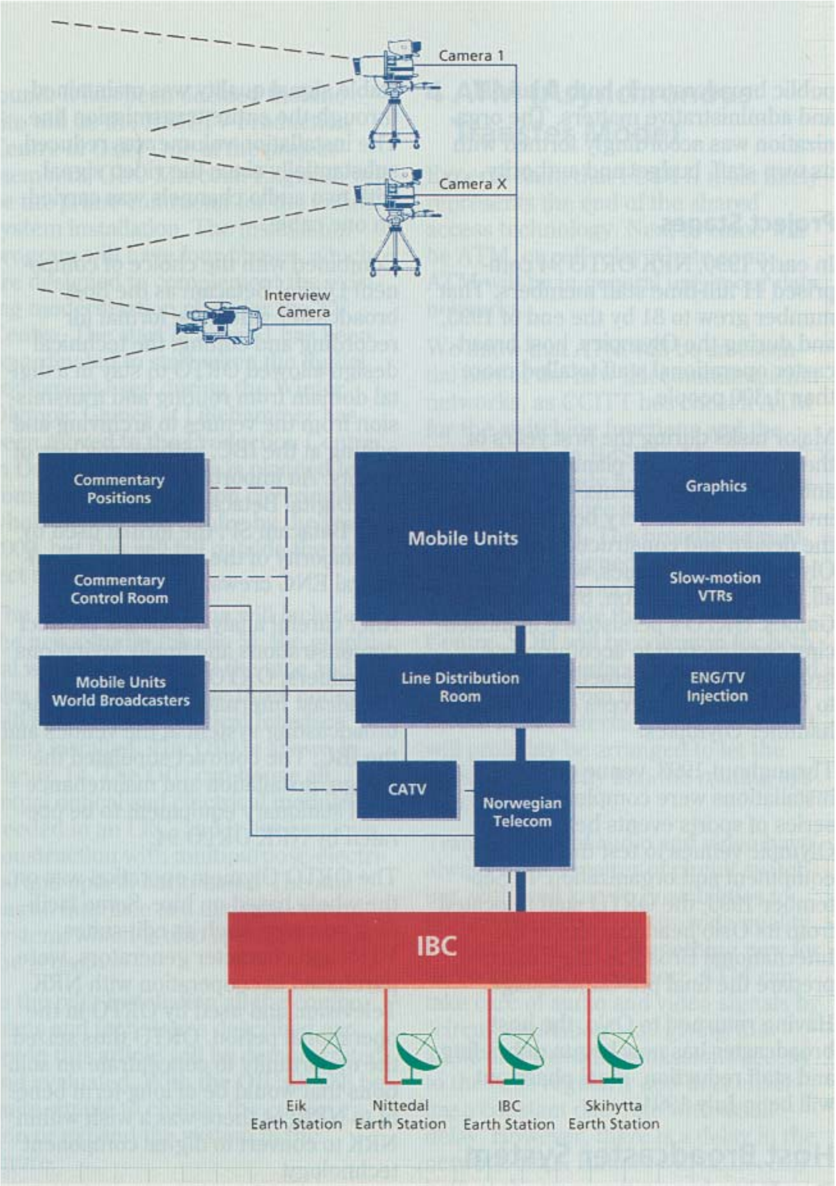


Figure 1 Venue Facilities Diagram

agreed to commit OB vans and production teams to the ORTO operation to cover those sports in which each broadcaster had demonstrated expertise. One-nation teams were appointed at each venue to guarantee fully cohesive coverage. Carrying the final responsibility, ORTO designed standard production and transmission guidelines for the international coverage, with the aim of conveying a distinct Lillehammer Games identity.

Under the aegis of NRK ORTO 94, NRK's own teams covered the opening and closing ceremonies, the cross-country skiing, the ski jumping and the Nordic combined events, SVT of Sweden covered the freestyle venue, YLE of Finland the biathlon, SRG of Switzerland and FT3 of France the two alpine venues, BBC of England the bobsleigh and luge, NOS of Netherlands the speed skating events and DR of Denmark the medal award ceremonies and press conferences. Canadian Television produced ice hockey, figure skating and short-track speed skating.

Each commissioned production team supplied its own mobile production equipment – 26 OB vans and more than 230 cameras in total. ORTO provided the stationary installations and also several special cameras for innovative and dynamic coverage, including sophisticated parallel-motion Cablecam and rail tracking cameras, mini cameras and helicopter and airship Wescam systems. Super slow-motion cameras were used at most venues.

In addition to three eagle eye cameras designed for the ORTO productions, four panoramic cameras were available at the IBC for broadcasters' unilateral use. Three of these cameras were remote controlled from the IBC and provided scenic views of the Olympic area 24 hours a day, whereas the fourth was a roving mobile unit operating at different locations.

Venue Operations

The great demands made on the quality of the Olympic coverage and the magnitude of the operations called for an extensive and carefully planned venue structure. The technical design (see Figure 1) was standardized at all venues, each with a Line Distribution Room or LDR serving as the local communications centre. All video and audio signals were handled in the LDR, while commentary and coordination circuits were managed in a separate Commentary Control Room, the CCR.

ORTO was responsible for all communications and signal distribution locally at the venues. Fibre optic systems as well as traditional cabling and RF equipment were used for this purpose.

At most Olympic venues, the ORTO production was carried out in analogue PAL. The video signal was digitized and decoded at the output of the production unit, and routed with graphic inserts to the LDR as a 270 Mbit/s signal with audio embedded. (See Figure 2.)

Commentary operations entailed 3.4, 7 and 15 kHz circuits with 3.4 kHz feedback, and 4-wire coordination circuits.

From the LDR and CCR the signals were passed on to Norwegian Telecom who was responsible for transmission to the International Broadcasting Centre. Video and audio from all venues were transmitted to the IBC on 140 Mbit/s fibre optic feeds. Telecom was responsible for the bit reduction and for reconverting the signal from 140 to 270 Mbit/s at the IBC.

Thus, there was a vanda (video and audio) signal in the true sense of the word, as opposed to previous Olympics, where video and audio were handled separately. These were also the first Olympics where the international sound was produced and delivered in stereo from all venues.

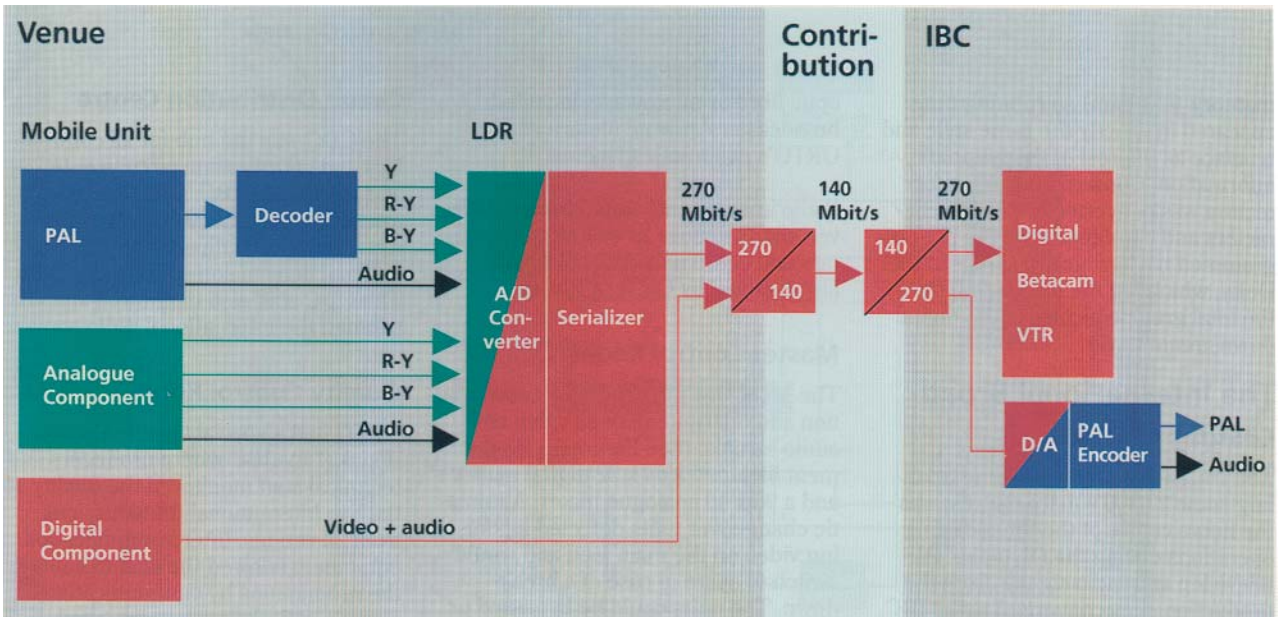


Figure 2 Transmission Principles Venues – IBC

The international radio signal was produced in analogue stereo in a separate mixer at the venue and transmitted digitally on Telecom's circuits to the IBC.

A special videotape facilities room was installed at each venue, equipped with Digital Betacam VTRs for local backup and clean feed recordings of the international signal.

ORTO offered a number of venue support services and facilities to broadcasters for their unilateral operations. At the eleven sports venues, ORTO installed and equipped a total of almost 570 commentary positions. Other unilateral services offered by ORTO included live interview cameras for occasional pre and post event booking, and ENG and TV injection points for transmission of taped programme from the venues to the IBC.

Backup Systems

All international and unilateral connections between the venues and the IBC were secured by Telecom's fibre ring structure with two separate transmission systems. For the international signal, a spare feed was routed permanently on a separate fibre so that there was a hot spare through the entire system.

Extensive measures were also taken against power supply breaks. Diesel generators were used as backup for the venue technical installations and as a further precaution, all facilities used in the production of the international signal had an uninterrupted power supply or UPS.

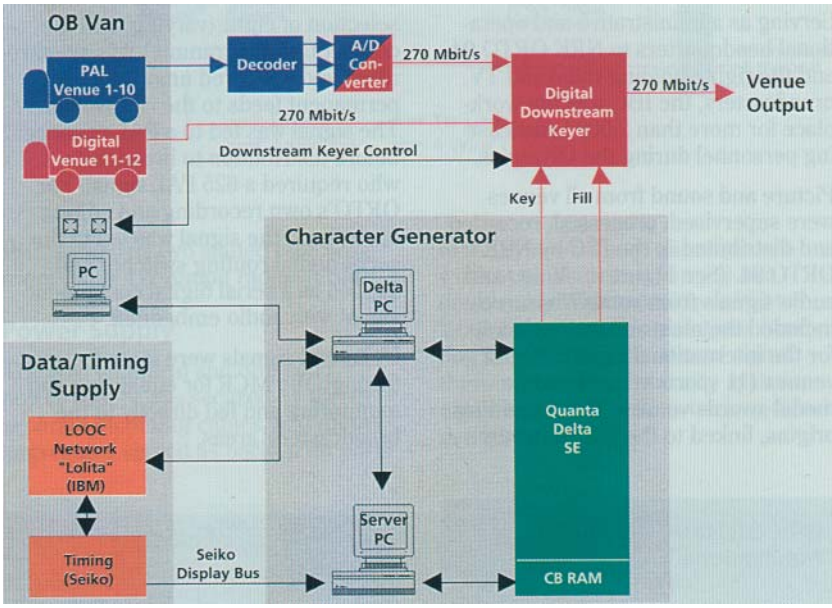


Figure 3 Graphic Inserts

Graphics

For the graphics operations, a standard structure was established at the venues to ensure an identical, high-quality graphic expression from all events (see Figure 3). The graphic standards had been pre-produced by ORTO designers in an all-digital suite. Since the graphics were superimposed in a digital domain, the inserts guaranteed the same output at all times and no local adjustments were required.

In a separate room at the venue, furnished with all the necessary graphics facilities, ORTO collected the feeds from the LOOC results computer centre (IBM) and the timing supplier (Seiko). The two sources were coordinated by ORTO personal computers and fed to the character generator Quanta Delta SE. The character generator was furnished with a high-speed memory (CB Ram) so running time appeared in exactly the same style and typeface as the rest of the graphic information. Consequently, all information was converted into video characters and combined with the pre-programmed background graphics in the Delta, which fed the key and fill for the total graphic picture to the digital downstream keyer.

The International Broadcasting Centre

The IBC, a five-storey building totalling some 27,300 square metres, was the nerve centre of the broadcasting operations during the Olympics. All television and radio signals from the production sites converged at the IBC, which was also the gateway for outgoing traffic to destinations worldwide.

Serving as administrative and operational headquarters to NRK ORTO 94 and the rights-holding radio and TV broadcasters, the IBC was the workplace for more than 3,000 broadcasting personnel during the Olympics.

Picture and sound from all venues were supervised, processed, recorded and distributed in the IBC by NRK ORTO 94. (See Figure 4.) Video and audio signals from some 75 sources included the main and spare circuits for the international signal from 12 venues (11 sports venues and the medal awards venue). The remaining origins, linked to the IBC either by optic fibre or microwave, included broadcasters' private feeds and ORTO's panoramic cameras.

Radio international sound from the eleven sports venues as well as commentary and coordination circuits were also handled by ORTO's IBC staff.

Master Control Room

The MCR was the centre for coordination and distribution of all video and audio signals. (See Figure 5.) Equipment included a 64 x 32 digital matrix and a 96 x 40 analogue mixer. Automatic changeover units detected incoming video on the main feed and would switch to spare in case of a breakdown. The units could be bypassed or forced to spare.

Among the main and spare international feeds from 12 venues, a daily selection of eight (varying with the competition programme) was synchronized and delivered unswitched on permanent feeds to the rights-holders. The signal was fed to a PAL encoder before distribution to broadcasters who required a 625 PAL signal. For ORTO's own recording and editing operations, the signal was fed to the serial digital routing switcher and applied as a serial digital component signal with audio embedded.

Unilateral signals were looped through the MCR for equalizing and monitoring and fed directly to the broadcasters' areas.

Circuit Distribution Centre

The CDC was the hub for all commentary and coordination circuits entering and leaving the IBC. ORTO received a total of roughly 1,500 circuits which were distributed directly on 250 international lines to some 60 destinations. The CDC staff also monitored and distributed the international radio sound.

Quality Control Room

ORTO had a special Quality Control Room at the IBC where producers and technical staff monitored the quality of the live international transmissions from all venues. Intercommunication equipment allowed the staff to control the productions by communicating directly with the venue personnel.

Transmission Control Room

During the Olympics, up to three competitions took place at the same time. The rights-holders made their own priorities as to which events to send live, and many of them supplemented the ORTO coverage – which was universal and strictly impartial – with interviews, comments, reports and other elements adapted to home audiences. Under final supervision in ORTO's Transmission Control Room, the output was patched to Telecom international circuits for transmission to the various destinations.

NRK ORTO 94 Olympic Publications

- Broadcasters Handbook
- Production and Graphic Standards
- Production Plan

General Statistics

ORTO personnel	1,360
World broadcaster personnel	4,200
Total hours of live ORTO coverage	350
TV audience reached	2 billion
ORTO OB vans	26
ORTO cameras	232
ORTO videotape recorders	200
Venue commentary positions	565
Venue unilateral camera positions	230

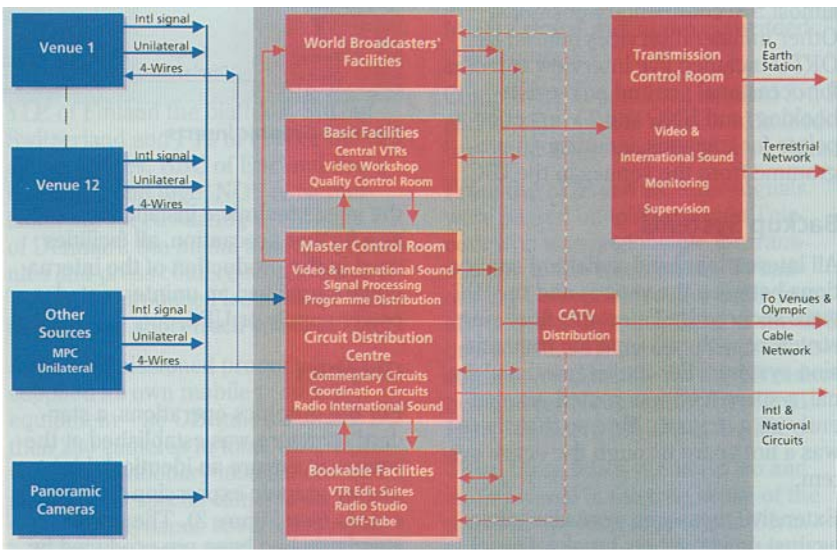


Figure 4 General Block Diagram Venues – IBC

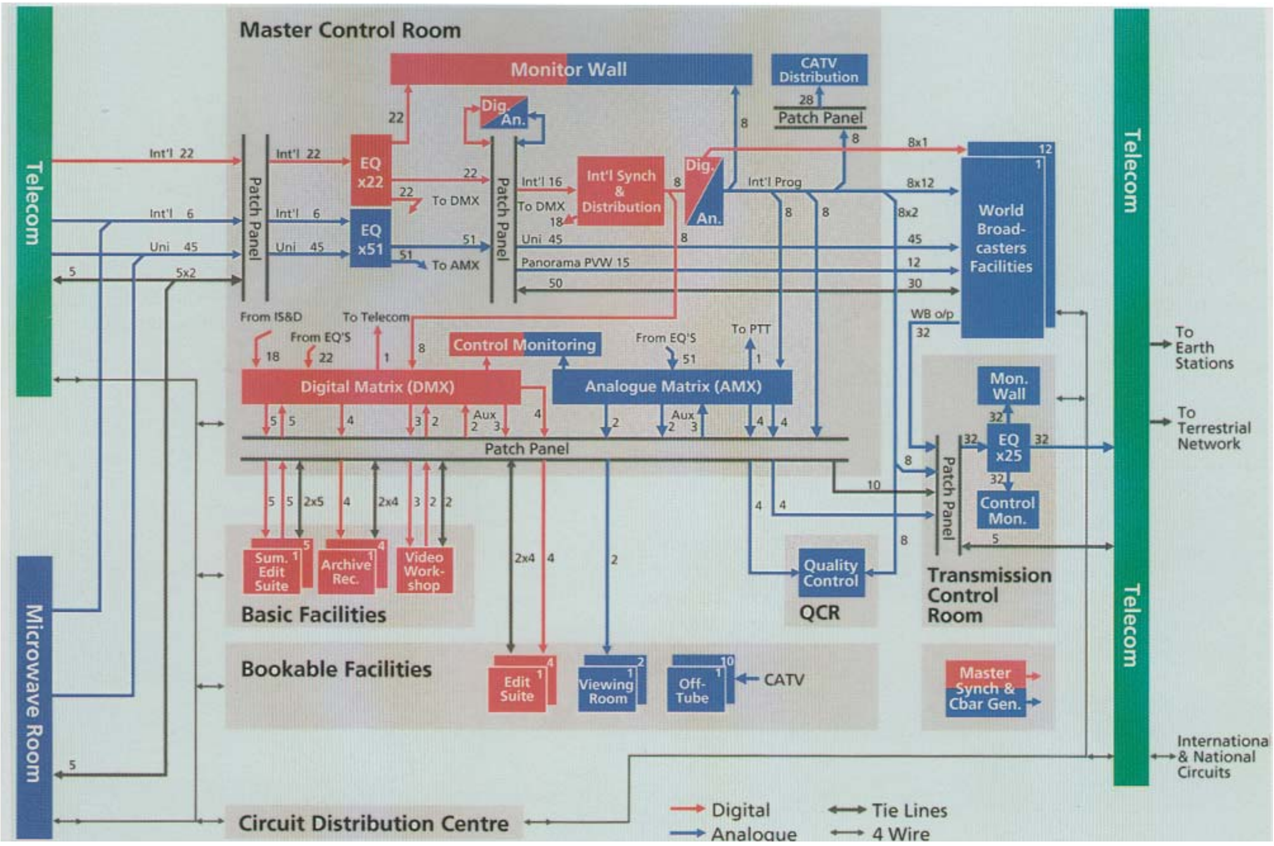


Figure 5 IBC General Flow Diagram

The Transmission Control personnel also operated a 48 x 8 router for switching of occasional satellite feeds for rights-holders.

Videotape Area

ORTO's central VTR area contained facilities for videotape recording and summary editing of the international signal. Four digital units equipped with Digital Betacam VTRs were designed for non-stop recording of the ORTO coverage from all venues.

Five edit suites based around Digital Betacam VTRs and Sony DVS-2000 vision mixers were used in the production of two daily 30-minute summaries presenting highlights from the Olympic events.

Bookable Facilities

A number of facilities were available to broadcasters for occasional use, including four staffed edit suites with three Digital Betacam VTRs and an BVW-D75 Betacam SP VTR each. This setup allowed users to edit and record digital or analogue footprint as required. Broadcasters could also rent two viewing units for videotape playback.

Sixteen commentary booths for off-site live commenting were offered at the IBC, as well as two staffed radio edit rooms with a common radio studio.

Power Supply

The IBC technical and domestic power supply was secured by three separate incoming cables, and diesel aggregates served as backup for parts of the system. There were four network stations (transformers) in a double ring with capacities of 50 to 2,500 KVA. For all installations handling the ORTO international signal, there was a special security system entailing UPS as well as backup diesel generators.

ORTO Venue Statistics

Event	Personnel	OB Vans	Cameras	VTRs	Commentary Positions
Opening/Closing					
Ceremonies and Ski Jumping	106	3	26	9	72
Freestyle Skiing	47	1	13	11	29
Cross-country	113	3	34	15	59
Biathlon	114	4	25	10	48
Downhill and Super G	133	4	23	19	56
Slalom and Giant Slalom	109	2	21	14	60
Bobsleigh and Luge	106	3	28	13	31
Ice Hockey (Håkon Hall)	52	1	14	11	59
Ice Hockey (Gjøvik Hall)	55	1	14	12	48
Figure Skating and Short-track	53	1	12	12	64
Speed Skating	58	1	15	9	38
Medal Award Ceremonies	21	1	5	5	-
Press Conferences	14	1	3	3	-

Status

International research and standardization activities in telecommunication

Editor: Endre Skolt

Introduction

BY ENDRE SKOLT

Introduction

Since the Status section on international research and standardization activities was launched last year, we have had contributions in all the telecommunication areas listed in Table 1. With reference to Table 1 the reader will for each of the telecommunication areas have access to information such as the main standard organizations and research programmes involved, technical study areas, achievements so far, what can be expected in the near future, major contributions, etc.

The Status section also contains presentations of important standardization bodies or research institutions. Up till now the Status section has had papers on the International Telecommunication Union (ITU), and the European Institute for Research and Strategic studies in Telecommunications (EURESCOM). This issue of Telektronikk contains a comprehensive coverage of ongoing work in Telecommunication and Information Networking Architecture-Consortium (TINA-C). TINA-C was formed last year and the overall objective is to define and validate a software architecture allowing provision of sophisticated telecommunication services. Information such as organizational structure, work plan, results achieved, and concluding remarks highlighting concerns raised, are included in the paper. The paper is written by Mr Tom Handegård.

With regard to the telecommunication areas, the following papers are included:

- Telecommunication Languages and Methods
- Message Handling Systems
- Security.

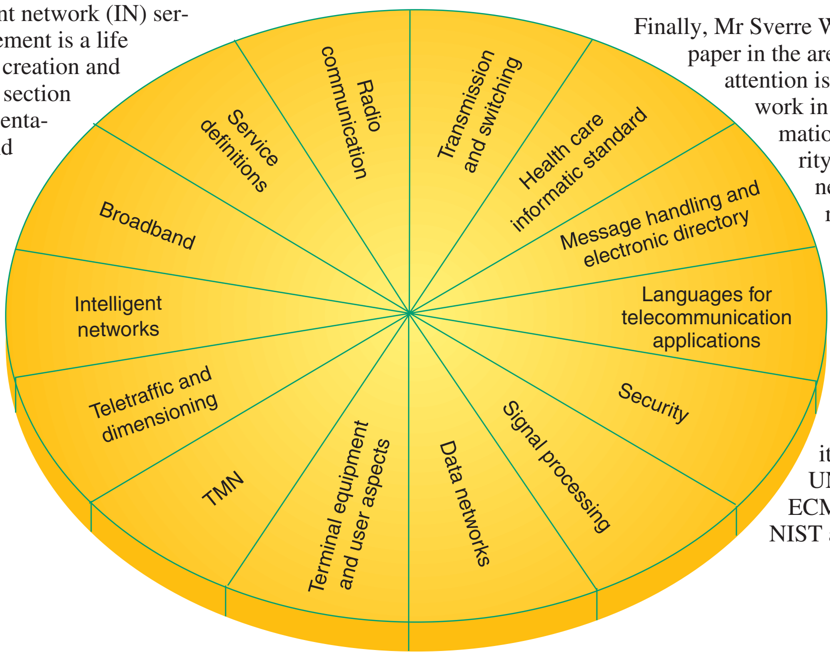
The paper on Telecommunication Languages and Methods presents activities and achievements in ITU-T, EURESCOM, and TINA-C related to this topic. Within ITU-T, SG 10 has the overall responsibility for co-ordination of Telecommunication Languages and Methods. In this study period SG 10 addresses 10 questions, which are briefly presented in this paper. EURESCOM P103 studies languages and methods for the creation of Intelligent network (IN) services. The main achievement is a life cycle model for service creation and composition. In the last section of the paper a brief presentation of the languages and methods being used in TINA-C is described and evaluated. TINA-C has adopted the object oriented paradigm and the concept of Open Distributed Processing (ODP). TINA-C also applies the viewpoints from ODP. The paper is written by Mr Arve Meisingset.

Table 1 List of contributions to the Status section

Issue No.	Study area	Editor
4.93	Service definitions	Ingvill H Foss
4.93	Radio communications	Ole D Svebak
4.93	Transmission and switching	Bernt Haram
4.93	Intelligent Networks	Endre Skolt
4.93	Broadband	Inge Svinnset
1.94	Terminal equipment and user aspects	Trond Ulseth
1.94	Signal processing	Gisle Bjøntegaard
1.94	Telecommunications Management Network (TMN)	Ståle Wolland
1.94	Teletraffic and dimensioning	Harald Pettersen
1.94	Data networks	Berit Svendsen
2.94	Languages for telecommunication applications	Arve Meisingset
2.94	Message Handling and Electronic Directory	Geir Thorud
2.94	Security	Sverre Walseth

Message Handling Systems (MHS) are an infrastructure for provision of application oriented data communication services. Today, the most important services are electronic mail, electronic data interchange (EDI) and file transfer. This paper introduces the development of X.400 standards, which is one of the most important manufacturer independent standards in the area of MHS. The author also presents the standardization groups and their area of responsibility, overview of the published standards and current and planned development of X.400. The author is Mr Geir Thorud.

Finally, Mr Sverre Walseth contributes a paper in the area of security. Special attention is given to the ongoing work in ISO/JTC/SC27 Information Technology – Security Techniques. In the near future important results are expected in security techniques such as entity authentication, non-repudiation, digital signatures, hash-functions and key management. Overall presentations of important work items in ITU-T, UN/EDIFACT, ETSI, ECMA, EURESCOM and NIST are also included.



The TINA Consortium

BY TOM HANDEGÅRD

Introduction

The TINA Consortium (TINA-C) was formed in 1993, and currently has as members more than 30 telecommunications and information technology companies from all over the world. The consortium will last for five years. Over this period the aim is to define and validate a software architecture that will enable the efficient introduction and management of new and sophisticated telecommunications services, such as broadband virtual network services, mobility services, and multimedia services. The name of the architecture is *Telecommunications Information Networking Architecture* – TINA. The *Information Network* is the TINA-C term on its vision of the future telecommunications network, a network capable of providing to its customers a wide range of services for accessing and manipulating information in many forms.

During its first year of existence, TINA-C has developed the first version of the TINA architecture. It is based on the principles of open distributed processing (ODP), and takes into account current standardization and research activities related to IN, TMN, SDH and ATM. The next version of the architecture will be published by the end of 1994.

Norwegian Telecom is a core member of the consortium, and contributes one researcher to the TINA-C Core team.

Background and vision

Several challenges are currently facing the telecommunications industry:

- It takes too long to introduce new services
- Services are too costly to introduce and maintain
- There are inconsistencies in the data needed to operate the network and its services
- The market for POTS is becoming saturated, and the telecommunications operators need to offer new services to increase their incomes
- New technology makes new services possible
- The number of actors in the business is increasing due to deregulation and new services, resulting in increasing competition.

Intelligent Networks (IN) and Telecommunications Management Network (TMN) are two current efforts set up to deal with different aspects of these problems, but not all of them. For example, current IN standards mainly focus on a set of interfaces and protocols which separate the service aspects from the basic call processing and switching of PSTN and ISDN networks. Current IN does not address the requirements of the upcoming broadband networks, to which more complicated call models apply. It is centred around the basic call service, and this will probably make it difficult to introduce truly different services. Furthermore, current IN and TMN standardization effort does not take advantage of recent advances in software design, such as distributed computing and object-oriented design, to help solve the problems connected to the service software's interoperability with management software, interoperability with software in increasingly powerful customer terminals, or the potential distributed nature of the service and management software itself.

To deal fully with these challenges the focus of attention must be moved from the physical network to the software system. There is a need for a software architecture that allows services and management applications to be introduced and maintained quickly and at a low cost. Issues like portability, interworking and reuse of software are of prime importance. Moreover, the architecture must support multiple suppliers, e.g. third-party development of software, and integration of heterogeneous technologies.

To deal with these software problems, one should exploit relevant results from the software engineering field. In particular, the field of *open distributed processing* (ODP) seems to address many of the issues mentioned above. Central to the ODP work is the vision of a "software bus" which supports interaction among software components in a similar way as a hardware bus support interaction among hardware components [14].

The goal of the software bus is to support pluggability of components conforming to the same interface, so that it is easy to plug one component out, and plug in another compatible component or one that implements an extended interface. The interoperability among components will be supported if they request and provide services through the mechanisms of the software bus. Existing applications (i.e. applications developed without the software bus in mind) can be plugged into the bus via an adaptor.

Many argue that the upcoming object-oriented technology is a promising tool for achieving this goal. Indeed, the majority of efforts within the ODP field applies object-based or object-oriented technology. When component interaction is limited to high-level inter-object messages, distribution can be made *transparent,* meaning that the objects are not able to tell whether they are communicating locally or remotely. This enables flexible placement of software modules. Integration of heterogeneous components is also facilitated when the only mode of interaction is via high-level messages, encapsulating object internals and system details. Furthermore, objects provide the level of granularity necessary for a workable approach to system reconfiguration supporting rapid changes and extensions, as well as a focus for reliability mechanisms such as object replication.

There are currently many efforts within the software engineering community devoted to development and standardization of architectures based on the software bus paradigm, notably

- ISO and ITU are working jointly on a *Reference Model of Open Distributed Processing* (RM-ODP) [13], currently having the status of draft international standard.
- The Object Management Group (OMG) has standardized a software bus, known as *CORBA* (Common Object Request Broker Architecture) [10]. Several software houses currently deliver CORBA conformant products.
- The Open Software Foundation (OSF) has defined and implemented an environment for distributed applications called

Distributed Computing Environment (DCE), currently being marketed by the individual members of the consortium.

- ANSA has developed an architecture that has provided significant input to the standards (RM-ODP and CORBA), and also developed a pilot implementation.

Figure 1 illustrates some principles connected to the software bus idea.

Telecommunications services and management could be viewed as software applications distributed over the nodes of a network, and communicating through a software bus. This is the essential idea in TINA's vision. TINA's name on the software bus and its associated components is *Distributed Processing Environment* (DPE). An overall illustration of the TINA architecture is shown in Figure 2.

Goals

The objectives of TINA-C are formulated in the consortium charter [4], and should be interpreted on the background of the ideas described above:

- *To define a TINA that is based on state-of-the-art distributed processing and service delivery technologies and that will enable efficient introduction, delivery and management of telecommunication services and infrastructures*
- *To validate the effectiveness of the architecture through laboratory experiments and field trials*
- *To promote the use of TINA world wide.*

The work will result in a set of specifications of the architecture, its component parts (including a library of general objects), and its key interfaces. Moreover, the consortium shall promote the use of TINA by running laboratory experiments to evaluate the architecture and prove that it can be implemented, to submit TINA specifications as proposals to standards bodies, to publish the results in the literature, and to participate in conferences and workshops.

On the other hand, TINA-C shall not make products. Also, hardware architectures of switching, transmission and computing equipment does not fall within the scope of TINA-C. The consortium members are not obliged to use the TINA-C results.

Guidelines for the work

A set of "guidelines" have been agreed for how the work to reach the goals should be conducted. First of all, the work is organized as a project. A team of resident researches, the so-called *Core team,* is set up in the consortium headquarters, currently located in Bellcore's facilities in Red Banks, New Jersey, USA. Each of the core member organizations supplies the Core team with one or more researchers. Currently, the core team consists of approximately 40 persons. This team of individuals working closely, concentrated, and full time on a single task, is considered one of the major strengths of TINA-C, and distinguishes TINA-C from other current research initiatives within the telecommunications field, such as RACE and EURESCOM.

Next, the strategy of the TINA work is to reuse existing de facto and de jure standards and other publicly available results as far as possible. Thus, the focus of the research is integration of existing concepts and technologies, rather than invention of new ones. In the first year of TINA-C, the Core team's work has particularly applied OSI/ITU and the Object Management Group's results within the field of open distributed processing, ISO and

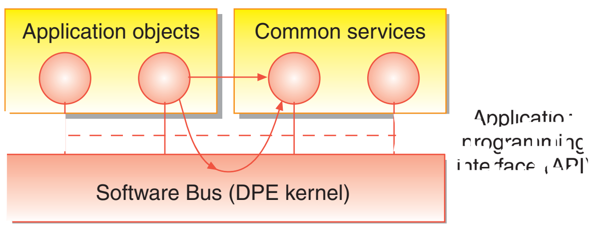


Figure 1 A simple software bus architecture
All objects communicate with each other through the software bus. The software bus is not directly visible to the application programmer, though. At the programming language level, objects communicate with each other only through inter-object operation calls (straight arrow in figure). Programming tools translate these calls into appropriate calls to the software bus API. Internally, the software bus may dynamically select one of several available communication protocol stacks to convey the operation call.
The software bus API and the rules for interaction between the software bus components on different computing nodes must be standardized. This will make the application objects portable, and enable interworking across different nodes, vendor products, and underlying technology (computer hardware, operating system and communication protocols).

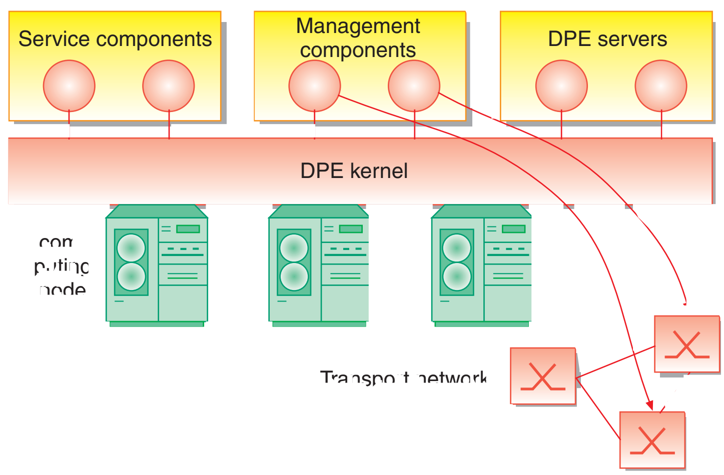


Figure 2 Overall TINA architecture
The software components are distributed across a network of computing nodes, and communicating through the DPE. Three categories of software components are identified: Service components, management components and DPE servers (general services extending the functionality provided by the DPE kernel). The interface between components managing network elements and the network elements themselves may be outside the scope of the DPE.

ITU's standards within Network Management and TMN, and Bellcore's Information Networking Architecture (INA).

Moreover, the TINA work shall have a unified view of services and management applications, and show how they are integrated within an ODP framework.

Table 1 TINA-C core members
eq = telecommunications equipment provider,
op = operator,
cmp = computer or software provider

Alcatel (eq)	NEC (eq)
AT&T (op, eq)	Norwegian Telecom (op)
British Telecom (op)	NTT (op)
Bellcore (op)	Nokia (eq)
CNET/France Telecom (op)	OKI (eq)
CSELT (op)	Royal PTT Nederland (op)
DEC (cmp)	SIP (op)
Deutsche Bundespost Telekom (op)	Stratus Computer (cmp)
Ericsson/Ellemtel (eq)	Stentor (op)
Fujitsu (eq, cmp)	TeleDanmark Research (op)
Hewlett-Packard (cmp)	Telefonica (op)
Hitachi (eq)	Telia (op)
KDD (op)	Telstra (op)
Korea Telecom (op)	Northern Telecom/BNR (eq)
Siemens (eq)	

Finally, the architecture will be validated and proven through laboratory experiments and auxiliary projects. The Core team will itself build a prototype DPE (again, reusing existing products), and design a small set of example services using the architecture principles. In addition, auxiliary projects are run within member companies' home locations. These projects will either test TINA, by building TINA conformant service prototypes, or investigate sub-problems the Core team regards as important, but does not have resources to deal with themselves.

Organization

There are two types of membership. *Core members* are organizations that contribute by sending one or more researchers to the Core team. *Participating members* have no representatives in the Core team, but instead perform one or more auxiliary projects at their home location. Core members carry the costs of their own visiting researchers. Additionally, all members share the administrative costs of the Consortium Headquarters, including the Core team. There is no membership fee in addition to this.

Member companies of TINA-C include telecommunications network operators, telecommunications service providers, telecommunications equipment manufacturers and computer manufacturers, software suppliers to the telecommunications industry, and research organizations. For a complete list of core members, see Table 1. In addition to the current members, several others are in the process of joining or have expressed their interest, such as Amdahl, EURESCOM, IBM, Samson, Swiss PTT and Versant.

The topmost level in the TINA-C management structure, illustrated in Figure 3, is the *Consortium Steering Board* (CSB). Only core member organizations that are network operators have a seat in the CSB. The CSB is responsible for evaluating the objectives and technical directions of the consortium, and amend them if necessary. The CSB reviews the technical activities annually, and all technical activities initiated within TINA-C must be approved by the CSB. The CSB also deals with membership questions and allocation of resources to the Core team and the administration office. Currently, the CSB is chaired by Mr. T. Rowbotham of British Telecom.

The *Consortium Technical Committee* (CTC) consists of one technical expert from each core and participating member organization. The CTC has the responsibility for the direction of the consortium's technical activities, and advice the CSB

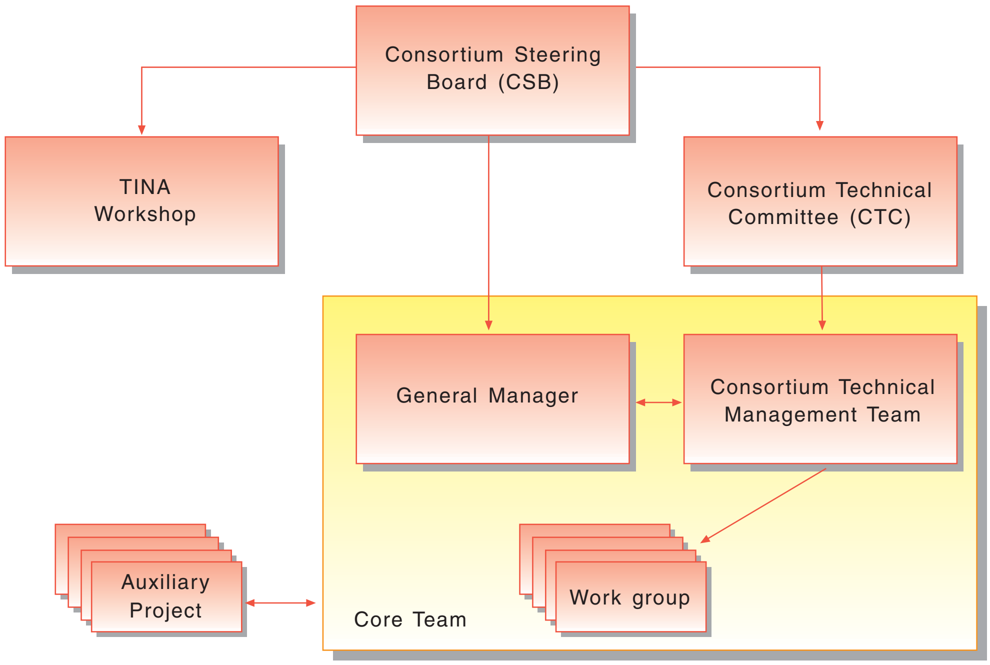


Figure 3 TINA management structure

on these matters. New technical activities are proposed by the CTC. The CTC is currently chaired by Mr Y Inoue of NTT.

Within the *Core team* the researchers are organized in *work groups;* currently there are four of them. The 1993 achievements of the work groups are described in the next section. *The Core team is led by the Consortium Technical Management Team,* chaired by the *Technical Leader.* A *General Manager* is responsible for managing and administering the day to day affairs of the consortium. Currently, Mr D Guha of Bellcore is General Manager of TINA-C.

The work plan time line

The TINA Consortium will last for five years, i.e. through 1997. A revised version of the architecture will be published by the Core team to the member companies at the end of each year. The main architectural results are expected to be finished by the end of 1995, with the last two years revisions being driven by results from the Auxiliary projects.

Results achieved in 1993

The 1993 work in the Core team was divided between four work groups, producing a total of 22 deliverables. An executive summary of the results can be found in [1], and a summary of all the reports in [2]. The availability of the TINA-C deliverables is currently restricted to consortium member companies only. This section aims to give a simple explanation of some of the main results.

The four 1993 work groups were:

- Logical Framework Architecture
- Service Specification, Construction, and Management
- Management Architecture and Resource Management
- Portable Reference Implementation of the TINA-C DPE.

I will refer to them simply as the Framework, Services, Management and Lab groups, respectively. The reports from the three first work groups define the first version of the TINA architecture. The Lab group mainly worked on the specification and design of a prototype implementation of the TINA-C DPE.

Framework

As already explained, a basic assumption in TINA-C is that services and management applications can be viewed as distributed software applications. The Framework defines the basic concepts and infrastructure (including DPE) needed to specify, implement, and run these applications. The Framework is based on ISO/ITU's RM-ODP, interpreting it where it is found too vague and expanding it where there are needs not covered. For example, RM-ODP does not specify any notations to write specifications. The TINA framework does.

According to RM-ODP, distributed systems can be described from different viewpoints, and RM-ODP identifies five such viewpoints. TINA-C has chosen to focus on three of them: The information viewpoint, the computational viewpoint and the engineering viewpoint. The first two focus on the structuring of the individual applications, while the engineering viewpoint focuses on the structure of the DPE.

Within the *information viewpoint,* an information model of the application is made. The purpose of the information model is to identify all the information entities, or information objects, that need to be managed by the application, the relationships between them, and the constraints that apply to objects and relationships. The TINA notation for documenting information models is based on ITU's *Guidelines for the Definition of Managed Objects* (GDMO) [11] and *General Relationship Model* (GRM) [12].

Within the *computational viewpoint,* a computational model of the application is developed. The computational model identifies the software modules that cooperate to perform the functions of the application. The modules are modelled as computational objects. In addition to the objects themselves, the computational model identifies (and defines) the objects' interfaces and their interaction structures. Thus, the computational model corresponds to the application designer and programmer's view of the application. The TINA notation for computational interfaces is an extension of CORBA's *Interface Definition Language* (IDL) [10].

The relationship between the information and computational models for a given application can roughly be described as follows. The information model contains objects that represent all the information entities relevant for the application, as well as their relationships, irrespective of how and by who these entities are managed. The computational objects are software objects actually existing in the system. Each of the computational objects will typically be responsible for managing a subset of the information objects. This may be reflected by the fact that the interface of a computational object contains operations to manipulate the information objects the computational object is responsible for.

The *engineering viewpoint* defines the internal structure of the DPE, and the mechanisms needed to deploy and run the implemented computational objects on the DPE.

Services

The central part of the work within the Services group is the definition of the so called *service architecture*. The logical framework described in the previous subsection is quite general, and could in principle apply to all types of distributed software applications. The service architecture expands the logical framework by defining principles and generic, reusable objects that are more specifically relevant and useful for telecommunications applications.

The most important elements of the Service Architecture are:

- The service lifecycle model
- The Universal Service Component Model (USCM)
- An initial set of reusable components, including a session model.

The service lifecycle model identifies all the stages a service goes through during its lifetime, beginning with the identification of the need for the service, through construction, deployment and operation, until finally the service is withdrawn. The lifecycle model serves to place the discussion of the various issues connected to services in a context.

The *USCM* identifies the separations that need to be adhered to when defining and implementing a service. The idea is that when a service is designed, some of the functionalities that are needed are so different (e.g. different paradigms and tools might apply) that they should be treated separately. The different separations, or types of functionalities, identified by USCM are functionalities needed to *access* the service (e.g. user interfaces), to *manage* the service or its components (management interfaces), to interact with the *environment* of the service (e.g. network resources), and finally the *core* logic of the service.

The set of *generic runtime components* identified partly constitute a *session model,* and partly cover service management issues. The objects provide functionalities such as mobility and independence of a service from specific terminal equipment. The identified components are likely to be needed and consequently reused by most services.

In addition to the service architecture, the services work group has also developed a *service design methodology.* The methodology is a step-by-step process that is meant to guide a service designer in the design of services so that they conform to the service architecture.

Finally, the services group has defined prose descriptions of a set of *example services,* the intention being that these services will be used to test the TINA architecture and exemplify the use of it. The example services chosen are Broadband Virtual Private Network, Nomadic Personal Communication (a variant of UPT), and Multi-media, Multi-party Conference.

The service design methodology and the example services are not normative parts of the TINA architecture.

Management

In the same way as the service architecture extends the general framework with definition of principles and reusable components relevant for the design of services, the management architecture specifies principles and reusable components relevant for management applications. The scope of management applications is control, monitoring and coordination of the network resources, both equipment and software, provided by a network operator. Equipment and software are likely to be supplied by multiple vendors. Important goals of the management architecture are to enable reuse of management software across multiple supplier's products and across multiple end-user services, as well as interworking of different management applications.

The management architecture is heavily reusing existing OSI management and TMN standards, but extends them where needed in order to allow consistent use of the distributed processing principles (as defined by the logical framework).

The TINA management architecture separates management into five functional areas: accounting management, connection management, fault management, performance management, resource configuration management, and security management. In 1993 specifications have been provided for three of these areas. Thus, the management architecture currently consists of the following parts:

- A generic network information model, consisting of managed objects available to all functional areas
- Specification of information and computational models for the functional areas connection management, fault management and resource configuration management.

The *network information model* incorporates a set of managed object classes from existing standards. In accordance with TMN principles, these managed object classes are abstractions of physical (e.g. switches) or logical (e.g. connections) network resources. The model also identifies relationships between the object classes. The object classes in the model are independent of transmission and switching technology, but are biased towards application in a broadband network (i.e. SDH and ATM) context. Detailed specifications, using GDMO and GRM are currently given only for a small subset of the object classes. The idea is that the object classes in the network information model should be available both to services and management applications, and thereby simplify design of new services and management applications.

The *connection management, resource configuration management,* and *fault management* definitions contain reusable functions and objects that are believed to be useful for a wide variety of services and management applications. In particular, most services are likely to be clients of connection management. Connection management is currently mainly focused on functions for establishing and releasing transport connections. The clients of connection management are allowed to manipulate connections in a technology independent way.

Lab

The long-term goal of the lab activity is to provide a prototype implementation of the TINA-C DPE. In 1993 the work has been focused on producing a specification of the DPE (i.e its application programming interface), and an implementation design. This work is directly driven by the specifications of the logical framework, especially the computational model (which defines the DPE's application programming interface), and the engineering model (which defines its internal structure).

The 1994 work plan

During its first year of existence, the Core team produced the first version of the TINA architecture. The output of the work was substantial: 22 deliverables, amounting to a total of more than 2000 pages.

During 1994, the architecture will be revised and refined. The second version of the architecture will be published by the end of the year. More specifically, the expected products of the 1994 effort are [5]:

- A revised and extended version of the architecture
- An object library, i.e. a set of specifications of reusable objects of the TINA architecture
- A (partial) DPE implementation
- Promotional and educational material, e.g. technical papers and presentations.

It has been agreed that the Core team during this year will be organized along the following four work groups [5]:

- Architecture extensions
- Architecture refinements through examples
- Laboratory work
- Marketecture (i.e. marketing of the architecture).

At the same time as working on the new tasks, the 1993 deliverables will be revised according to feedback from the 1994 work. Thus, a sort of matrix organization appears, with the 1993 deliverables along one axis, and the new work groups along the other. Not all 1993 deliverables necessarily need to be revised, and new deliverables may be identified.

The tasks of the new work groups are as follows. The *extensions group* will perform the same type of work done in 1993, but cover important issues not yet addressed. Four high-priority and six low-priority issues have been identified. The high-priority issues are: *Service and communication session management* (developing the model for service sessions that was sketched by the 1993 services group), *Connection management* (extend the existing model with multi-point connections and routing aspects), *Performance* (timing consideration, replication of software objects, reconfiguration of components in the network), and *Packaging* (define and integrate the various mechanisms related to the building of software modules). The low-priority issues are: Security, Naming, Federation, Service interaction, and Data management. The Core team will only do rudimentary studies on the low-priority issues, and then invite member companies to suggest solutions (e.g. through auxiliary projects).

The *refinements group* will design a set of example services. One of the problems with the first version of the architecture, is that different parts of it were developed by different persons working in parallel. The result is that the different parts are not sufficiently well integrated. The idea is that designing a set of services, using all parts of the architecture, will provide valuable input on how to improve the architecture in this respect. It is also expected that the design work will result in a collection of object specifications that will be incorporated in the TINA-C object library. The selected example services are: Broadband Virtual Private Network, Freephone, and Joint Document editing. The rationale for selecting these services are that they should demonstrate the usability of the TINA architecture in building, respectively, a set of management functionalities, an intelligent network service, and an advanced service in line with TINA's information network vision mentioned in the introduction of this paper.

The *lab group* will continue its work on designing and constructing a prototype implementation of the TINA-C DPE and other selected components of the TINA architecture.

The mission of the *marketecture group* is to coordinate the promotion (or marketing) of the architecture, by publicizing TINA-C results in member companies, conferences, workshops, standards bodies and the industry news media. This activity also includes creation of documentation and training material that is targeted to different audiences, such as systems analysts, standards organizations, other consortia, users of telecommunications services, and application developers.

Current auxiliary projects

As previously mentioned, auxiliary projects are carried out in the home locations of consortium member companies. The purpose is to supplement the effort of the Core team. The auxiliary projects are classified within the following categories [8]:

- Projects which are *customers* for the entire architecture. Such projects take the complete architecture, build experimental applications, and feed the results back to the Core team.
- Projects which are *contributors* to the architecture. Such projects are typically initiated by a Core team request for solution (as with the low-priority issues of the extensions group described in the previous section).
- Projects which independently *asset* the architecture using criteria like time to market, availability of reliable technology, or evolution scenarios.

Proposals for new auxiliary projects are submitted to the CTC, and have to be approved formally by the CSB. Once approved, one member of the core team is selected as "champion" of the project. The job of the champion is to represent the interests of the auxiliary project in the Core team.

Currently, nine auxiliary projects are approved. They are:

- Platform for TINA Prototyping (Telstra), uses UPT as an example application to develop reusable software components
- Personal Presence System (Bellcore), investigates TINA-consistent management of a multiparty, multimedia bridge
- Data Management Platform (CNET), investigates issues of data management within a TINA-C DPE environment
- Service Management Experiment (Korea Telecom), validates the TINA-C service architecture by using UPT and VPN as example services
- Remote Data Access and Transaction Services in TINA (Unisys, Wollongong University), investigates the use of RDA for transactions
- Value Analysis of the TINA-C Architecture (Stentor, Stratus, Northern Telecom, Bellcore, Hewlett-Packard), describes the business benefits that can be derived from using the TINA-C architecture

- Practical TINA-C Architecture Validation (British Telecom), validates the TINA-C architecture by using a multimedia application as example service
- DPE on CHORUS (CNET), aims to demonstrate the TINA-C DPE on the CHORUS real-time operating system
- Personal Communications Service (Bellcore), investigates how to provide a highly reliable Personal Messenger Service on the TINA-C DPE.

Concluding remarks

This paper has given an overview of the current status of the TINA Consortium. During its first year of existence, TINA-C has made tangible progress towards its goal of defining an open architecture for telecommunications applications based on ODP principles. Considering the technical complexity of the subject, it should be no surprise that there are also sources of concern, areas that perhaps are not sufficiently addressed by TINA (some of them are addressed in the 1994 work plans, though). First, the architecture work can be accused of putting to much emphasis on problems which are traditionally associated with software engineering, sacrificing problems of crucial importance to telecommunications, e.g. realtime, reliability and security requirements.

Moreover, the TINA architecture makes almost no reference to current IN. TINA uses object-orientation, has introduced a totally different connection management scheme, and is generally incompatible with the IN Conceptual Model. (This is in contrast to the management area, where TINA does not make this total paradigm shift, and has been able to reuse existing standards to a much larger extent.) As mentioned early in the paper, TINA is primarily targeted at new services and new (broadband) networks. The paradigm shift within the services field is done because it is considered necessary. However, some argue that the migration from services based on existing architectures towards TINA should be addressed.

The decision of the Core team not to focus on a service creation environment also concerns some. Certainly, the object library highlighted as one of the main results of the work will be an important element of the service creation environment. Clearly, an object library alone is not sufficient, though.

The last area of concern to be mentioned, is that the process of influencing standardization with TINA results has not started, nor is there currently a strategy for contributions to relevant standards bodies. One of the first relevant areas could be connection management. TINA connection management specifications could influence (and be influenced by) ITU standards on B-ISDN.

Anyway, the current TINA results are preliminary, and should be judged in the light of the fact that TINA-C has only existed for about a year. Bringing researchers with different kinds of expertise, and from different companies, together in a Core team has proven successful. It contributes to the development of common understanding of important issues, and gives a good utilization of scarce resources. Even though all parts of the TINA-C work not necessarily will become equally important, some parts will *most likely* become important, and to a certain extent already are. For example, the TINA principles of using an ODP framework for building telecommunications software is currently being promoted and applied in several projects within EURESCOM and RACE.

References

1 *An overview of the TINA Consortium Work,* Document No. TB_G1.HR.001_1.0_93, TINA-C, December 1993.

2 *TINA-C Documentation Roadmap,* Document No. TB_G.RM.003_1.0_93, TINA-C, December 1993.

3 *Press Release Announcing the Formation of the TINA Consortium,* to be published.

4 *TINA International Consortium Charter,* TINA-C, October 1992.

5 *TINA-C 1994 workplan,* TINA-C, January 1994.

6 Report on CTC Audit of Core Team Activities, TINA-C, January 1994.

7 *TINA-C Newsletter,* 1(1), TINA-C, April 1994.

8 Auxiliary Projects – A Framework for Cooperation, TINA-C, January 1994

9 Barr W J et al. The TINA Initiative. *IEEE Communications Magazine,* 31(3): 70–76, 1993.

10 *The Common Object Request Broker: Architecture and Specification,* OMG, December 1993.

11 ITU Recommendation X.722, *Information Technology – Open Systems Interconnection – Structure of Management Information – Part 4: Guidelines for the Definition of Managed Objects (GDMO),* September 1991.

12 ITU Recommendation X.725, *Information Technology – Open Systems Interconnection – Structure of Management Information – Part 7: General Relationship Model,* January 1993.

13 ITU Recommendation X.903, *Basic Reference Model of Open Distributed Processing – Part 3 – Prescriptive Model,* June 1993.

14 Berre, A J, Handegård, T and Løvnes, K. *Open Distributed Processing and IN* (preliminary title), unpublished draft.

Telecommunications languages and methods

BY ARVE MEISINGSET

Norwegian Telecom Research currently participates in international activities on telecommunications languages and methods in the following fora:

- International Telecommunication Union –Telecommunication Standardization Sector (ITU-T) Study Group 10 (SG10) Languages
- EURESCOM Project 103 (EU-P103)
- Telecommunications Information Networking Architecture Consortium (TINA-C).

ITU-T is the standardization sector of ITU. Eurescom is a joint research organization for European network operators. TINA-C is a consortium of some of the world's leading network operators developing a common platform for intelligent network applications. Some other fora will be mentioned in relation to the above. The Norwegian Telecom Network Division and IT department are involved in some of these.

Why are software languages important? When someone specifies or develops software systems, what they specify, the quality of the design etc. are highly dependent on the languages used. Often the impact of the language constructs used is unknown to the language user. Note that even informal specifications using natural languages are dependent on language notions used informally by the specifier, and the structure of the (informal) specifications has consequences for the design and functioning of the final system. Therefore, an understanding is needed of how the language in a broad sense influences the final design of the system. To some extent we are talking about expressive power and fine-grained features of the language; what is difficult to express and what is not. However, just as important are large scale issues; how the language contributes to layering, encapsulation, reuse or joint use of modules, what is specified and what is not, what is focused and what is not etc. This way, language issues are central to the design of future software based systems of any telecommunication vendor, operator or user. The design and operation of large real-time systems have been special to telecommunications. Now the telecommunication technology has changed to utilize large real-time communicating databases. The languages have to be chosen and adapted to the needs of this new technology and services.

The subsequent presentation will describe status of work and results from the three fora mentioned above and highlight the Norwegian involvement in this work. The work is presented in the mentioned sequence, as the ITU languages are used and referenced in both the succeeding projects.

ITU-T SG10 – Languages and methods in ITU

SG10 is in this Study Period addressing 10 Questions. Each Question has a rapporteur who together with the chairman and vice-chairman constitute the SG10 management board. The chairman is Ove Færgemand Eurescom/Tele Denmark and the vice-chairman is Klaus Schulz Deutsche Bundespost Telecom. SG10 co-operates with other SGs through the Telecommunications Standardization Advisory Group (TSAG), where all SG chairmen participate. Also, SG10 has been participating in the Joint Coordination Group for Telecommunication Management Networks (JCG for TMN). Finally, SG10 has direct co-ordination meetings with several SGs. Co-ordination with the International Standardization Organization (ISO) is undertaken through SG7. In addition, SG7 has adopted and developed some languages (GDMO and ASN.1) and methods from other bodies and submitted these to other SGs. Finally, formal and informal specification techniques for services (SG2), intelligent networks (SG11), telecommunication networks (SG4) and other have contributed to the diversification of languages and methods within ITU. This provides both a challenge and opportunity to the development and co-ordination work undertaken by SG10.

Arve Meisingset is head of the Norwegian delegation to SG10.

Question 1/10 – A data definition language for user data

Q1/10 – Improved methodology to specify Human-Machine Interfaces (HMI) – develops a formalism for the definition of a common user terminology and grammar for a large application area. Arve Meisingset, Norwegian Telecom Research, is the rapporteur for this Question.

A draft Recommendation (draft Z.353) exists for this Question. The proposed formalism has similarities to other data definitions languages, such as the Entity-Relationship formalism. However, there are important differences needed to define the end-user terminology as it actually appears on the screen. These features include the handling of local names (context sensitivity), significant duplicates, persistence of forms etc. The centralization of the definitions in a common application schema ensures harmonization across large application areas. Contributions exist on how to define references between data and some behaviour aspects, as well as the definition of and mappings between different alphanumeric and graphical terminologies. The Question addresses linguistic (deep structure) aspects of the HMI, not being addressed by other HMI standardization bodies.

Question 2/10 – A reference model for Human-Machine Interfaces

Q2/10 – Specification techniques for the presentation and manipulation of data at the Human-Machine Interface – develops a specification technique for the external presentation and manipulation of data on the HMI. Alan Burger, Siemens, is the rapporteur for this Question.

Two Recommendations are approved (Z.351 and Z.352). These define a reference model, approach, method, scope and terminology for HMI. The reference model is based on the 3-schema architecture for data bases. However, its use and actual contents are different. The strong data orientation makes the approach different from more commonly applied task oriented models for HMI. Z.352 includes an important Appendix providing guidelines for good data design.

The external layer of the HMI reference model comprises two sub-layers containing contents and layout schemata. These will be addressed in the near future. Also, contributions exist on how to design icons and scematic graphs to the end user, e.g. of the telecommunications network.

Question 3/10 – Application to TMN

Q3/10 – Specification of Human-Machine Interfaces to support management of telecommunication networks. This Question studies the application of the HMI Recommendations on Telecommunications Management Networks (TMN) applications. Currently the Question has no rapporteur. Therefore, the work is carried out under Q1 and Q2/10.

The Z.300 series Recommendations (Blue Book) contain Recommendations for the CCITT Man-Machine Language (MML). This includes the task-oriented SOF (Specification Of Function semantics) method. Experienced difficulties with this method caused the development of the new data oriented specification technique to be provided in the Z.350 series Recommendations.

The main application area of the HMI specification technique is TMN. Currently Recommendations exist for the data exchange between TMN systems and Network Elements (NEs) over the Q3 interface. The Recommendations include data definitions for this exchange. The data definitions use Guidelines for the Definition of Managed Objects (GDMO) and Abstract Syntax Notation 1 (ASN.1). Currently there is hardly any contents of the Recommendations for the exchange of data between TMNs (over the X interface) and between TMN and the outside world (over the F interface). GDMO is an object oriented language at about the same abstraction level as SQL for data bases. However, GDMO and ASN.1 will be found more technical than SQL by most users.

Q3/10

- is a testbed for making TMN Recommendations useful to the final users (over the G interface);

  Q3/10 has applied the HMI technique to Access Control Administration and Alarm Surveillance and found difficulties with both these TMN Recommendations

- is contributing to define the borderline between TMN and the outside world;

  Q3/10 has developed a draft interoperability reference model as an extension to the 3-schema architecture and given contributions to the redefinition of the F interface

- developed a graphical notation for GDMO to make TMN Recommendations comprehensible, manageable and mapable to the HMI.

Several other studies on HMI were undertaken in the previous Study Period which will not be presented here. The HMI work (G) currently suffers from small resources, similar to the situation for the work on the F and X interfaces (in SG4) of the TMN.

Question 4/10 – ITU adopts software quality standards

Q4/10 – Software quality for telecommunication systems – collects and reviews existing standards' documents on software quality. Approved results are found in Recommendation Z.400 – Structure and Format of Quality Manual for Telecommunication Software. A software quality life-cycle model and software quality criteria are expected short term results from this work. Yossi Winograd, AT&T, is rapporteur for this Question.

Question 5/10 – Software architectures

Q5/10 – Software architectures and platforms for distributed systems in the telecom domain. This question is currently in a dormant state as no resources are available to carry out this work. Related work is currently undertaken in Eurescom, Tina-C and Spirit (of Network Management Forum).

Question 6/10 – The ITU object oriented process specification language

Q6/10 – Maintenance and support of SDL. SDL is the specification and description language for telecommunication systems and has been used by telecommunication vendors and purchasers for a long time. The language is based on the notion of an extended finite state machine and asynchronous signalling, and has a graphical notation which has been found convenient for most users. Armadeo Sarma, Deutche Bundespost, is the rapporteur for this Question.

SDL has evolved to become a rich specification language and is documented in the Z.100 series Recommendations. The latest extension has been to include object oriented features in what is known as SDL'92. Birger Møller Pedersen Norwegian Computing Centre has been a central contributor. The object oriented features of SDL'92 have its root in the Beta and Simula languages from Norway. Rolf Bræk, Delab, has in the past given important contributions to the SDL method.

SDL has got an advanced data definition part, known as SDL Abstract Data Types (ADT). ADT has got a common kernel with LOTOS, standardized by ISO. While LOTOS has its root in the Communication Calculus Language (CCL), ADT stems from the algebraic specification language ACT ONE.

ADT has been found too academic and hard to use by most SDL users. This has led to investigations into integrating GDMO and ASN.1 into SDL. A GDMO Managed Object Class may be defined to be an SDL process. Another reason for this integration attempt is that GDMO and ASN.1 have got widespread use within and outside ITU, and SG10 has to direct its work towards the current market. Finally, GDMO needs a means to define the behaviour of Managed Objects; SDL is the ITU choice for this. In a longer time frame, other options will be investigated. Telecommunication Management Networks (TMN) and Intelligent Networks (IN) are considered being important new markets for SDL, and co-operation with the relevant SGs and adjustments to their needs are under way.

Q6/10 is currently the largest and most active group within SG10. The set of associate rapporteurs indicates the width: Birger Møller-Pedersen Norway for the Simplification of SDL, Rick Reed UK for Methodology, Lois Verard Sweden for SDL & ASN.1, Joachim Fischer Germany for Formal definition of the static part, Eva Hedman Sweden for Common Interchange Format, Armadeo Sarma Germany as liaison rappoteur to SG11, Rick Reed as liaison rapporteur to SG7, and finally Joachim Fischer Germany as liaison rapporteur to SG7 on Open Distributed Processing (ODP).

Birger Møller-Pedersen and Astrid Nyeng Norwegian Telecom Research have investigated the use of either SDL'92 or the HMI specification technique to the intelligent network service Universal Access Number (UAN). The conclusion of this study is that SDL'92 can be conveniently applied for the formalization of the current IN Recommendations, while the HMI specification technique can be a more appropriate candidate for the definition of future data intensive IN services. The study results are contributed to ITU and documented in research reports and papers.

SDL Forum is a yearly event for presentation of the SDL language, tools, extensions and its use.

Question 7/10 – Modelling techniques

Q7/10 – Modelling techniques for telecommunication systems – intended to define core modelling techniques for telecommunication applications, define suitable specializations for these applications and guidelines for their use. The Question intended to address new needs within ITU, in particular needs from Open Distributed Processing (ODP). However, currently attempts are made to adjust SDL to these needs and, therefore, Q7/10 is dormant.

Question 8/10 – Testing and verification

Q8/10 – Testing based on formal specification and verification of formal specifications – co-operates with ISO Project 54 in developing Recommendations for software testing and verification. Dieter Hogrefe, Switzerland, is the rapporteur for this Question.

Question 9/10 – Specification of message sequencing

Q9/10 – Message sequence charts (MSC) syntax and semantics. Z.120 provides a Recommendation for how to specify sequences of messages, which is a need for anyone specifying interaction between communicating systems. Ekkart Rudolph, Siemens, is the rapporteur for this Question.

In this Study Period the technique will be extended with structuring features. Øystein Haugen, Norway, is a key contributor and associative rapporteur of this work.

Question 10/10 – The ITU programming language

Q10/10 – Maintenance and evolution of CHILL. CHILL is the ITU high level programming language. This is a language in the Pascal and Algol tradition for the programming of real-time systems. Jurgen F.H. Winkler Jena University will be the rapporteur for this Question.

The CHILL language has been used extensively by telecom vendors and operators for many years. Kristen Rekdal, Norway, has been a key contributor to this language, and the Norwegian company Kvatro A/S –where Rekdal is marketing director – provides the only existing commercial compiler and tool support for the CHILL language, including SDL to CHILL and ASN.1 to CHILL translators. Several large telecom vendors provide their own tools.

After several years of maintenance, work is now under way to develop an object oriented CHILL. Dag H. Wanvik, Kvatro A/S, is the Norwegian contributor to this work.

EURESCOM guidelines for service creation and composition

Eurescom P103 studies languages and methods for the creation of intelligent network (IN) services. IN services require flexible definition and management of services according to user needs. This puts strong demands on service creation and composition, and the communication of these data to and from the telecommunication network. Raymond Nilsen Norwegian Telecom Research is the project leader of P103.

P103 is currently in its second phase (July'93–Dec'94). The first phase studied service creation, requirements on IN architectures and status quo studies on existing architectures.

P103 is developing a method – lifecycle model – for service creation and composition. This method comprises three major activities for analysing and designing a service:

i) dividing it into a number of aspects that can be treated independently

ii) elaborate each aspect into a role model, arriving at a complete specification with behaviour

iii) compose the individual specifications to a complete service specification.

The descriptions are made in OOram, MSC and SDL'92. The Norwegian contributor to the method, being developed in phase two, is Bengt Jensen, Norwegian Telecom Research. The OOram method stems from the OORASS method developed at Center for Industrial Research (SI) in Norway. This method and its tools are currently commercialized through the Norwegian company Taskon A/S. The method comprises Role modelling, Object specification, Class implementation, Structure configuration and System installation. MSC is described in the previous section ITU-T SG10 Question 9/10. SDL is described in the previous section ITU-T SG10 Question 6/10.

Additionally, the second phase will study architectural support for service provisioning. One initiative that will be evaluated is TINA-C; see the next section.

Finally, the chosen techniques will be applied to a test case to uncover possible problems.

Telecommunications Information Networking Architecture Consortium

The Telecommunications Information Networking Architecture Consortium (TINA-C) develops a common architecture for a multi-vendor telecommunication service platform. The architecture is designed to meet the following requirements:

- interoperability
- reusability of application software
- distributed processing and data
- support of new types of services
- support of management
- independence from computing environment and hardware
- quality of service
- scaleability
- security
- compatibility with existing technologies
- flexibility against regulation
- conformance testing.

TINA-C has been working for one year and the current presentation is based on deliverables of December 1993. Several open issues will be studied in 1994, and the architecture will be finalized by then end of 1997.

TINA-C has adopted the following as fundamental technologies to meet the requirements:

- object oriented paradigm (derived from ODP)
- distributed processing environment (DPE).

TINA-C applies the viewpoint notion from Open Distributed Processing (ODP) in ISO. The best defined viewpoints in TINA-C are:

- information viewpoint
- computational viewpoint
- engineering viewpoint.

The remaining viewpoints are the enterprise and technology viewpoints. TINA-C applies a transformational approach, as viewpoints of one development phase are transformed to that of a subsequent phase including additional details.

Currently, TINA-C prescribes no method. For comparison, we mention that Norwegian Telecom has experience with applying the transformational method SSADM, which is a governmental method in UK. Also, Norwegian Telecom Research has developed a transformational tool from SDL to CHILL and C.

The ODP viewpoints stem from the Advanced Network Systems Architecture (ANSA) and seem to have a bearing on the more rigorous Zachman framework.

Information modelling concepts

The information specifications provide the necessary knowledge to interact appropriately with a system. Internal aspects of the system are normally not dealt with in the information viewpoint.

TINA-C information viewpoint does not seem to have a reference model for the target system. Consequently, both ‘external’ aspects, like screen and reports, and ‘conceptual’ aspects, the data, are covered by the information viewpoint.

TINA-C informally defines and discusses a set of modelling concepts for the information viewpoint. However, TINA-C does not define its own formal language for these concepts. Rather it applies several existing formalisms and combines these to meet the TINA-C requirements. These formalisms have to be reinterpreted to fit the TINA-C modelling concepts, and TINA-C offers an evaluation and comparison of the different formalisms. The candidates are Guidelines for the Definition of Managed Objects (GDMO) and General Relationship Model (GRM) from ISO, OMT from General Electric – which has a convenient graphical notation for objects –, Object-Z – which originates from UK, and which is currently being standardized by ISO –, SDL from ITU, and COOLish (together with ROOM) from the RACE Open Service Architecture (ROSA). GDMO+GRM is recommended as the pragmatic choice, while Object-Z is the ideal academic choice. A problem with the use of GDMO+GRM is that it currently applies for the Telecommunication Management Network only, and not for the Telecommunication Network itself and the IN. Hence, interaction with the objects inside the telecom network – and IN – can bypass the GDMO+GRM specifications.

TINA-C recognizes that the information viewpoint is poorly developed in ODP. The ‘naming policy’ of the TINA-C information viewpoint is for further study.

The Norwegian contributor to TINA-C Information Modelling Concepts is Erik Colban, Norwegian Telecom Research.

Computational modelling concepts

The computational specifications specify the component design of the system. The computational objects interact with computational objects only. Therefore, information objects are computed (or represented), but not accessed, by computational objects.

A central theme in the design of the computational modelling concepts is to achieve distribution transparency. The computational viewpoint takes the Reference Model for Open Distributed Processing (RM-ODP) from ISO as its baseline document. Transactional operations are required to satisfy ACID (atomicity, consistency, isolation and durability) properties. Additionally, the computational viewpoint applies the Interface Definition Language (IDL) from the Object Management Group (OMG).

The computational viewpoint has borrowed several concepts from the Information Networking Architecture (INA) and the later OSCA architecture (e.g. building blocks and contracts), both from Bellcore. Also there are relations to the CSELT (Italy) Computational Model, SERENITE Computational Model – which cover multimedia services –, and the RACE Open Service Architecture (ROSA) – in particular the COOLish language.

Contrary to the information viewpoint, the computational viewpoint applies the OSCA functional separation rules, which state that functions of a building block should belong to either the human-user interface, the management of shared persistent data or non of these. This corresponds to the separation of the 3-schema architecture.

One object of the computational viewpoint can have several interfaces: be client and/or server to many objects, be producer or consumer of stream interfaces, and/or have both service and management interfaces.

TINA-C recommends to start with computational object definitions and derive interfaces from these definitions rather than the other way around.

Several open issues remain for the computational viewpoint; examples are timing constraints, and how queries can be expressed for data intensive applications. Finally, complex transaction processing and selective use of ACID properties are challenging areas to be studied for the TINA-C environment.

Engineering modelling concepts

The engineering infrastructure is an abstract model for the Distributed Processing Environment (DPE) and describes how distribution transparency is accomplished.

The engineering modelling concepts are based on the Reference Model for Open Distributed Processing (RM-ODP) in ISO and the Advanced Network Systems Architecture (ANSA) from Bellcore. The concepts are influenced by the Information Networking Architecture (INA), SERENITE, OMG/COBRA and CHORUS.

The runtime components of the DPE are the DPE kernel and a collection of DPE servers. The engineering modelling concepts comprises deployment concepts, communication concepts, systems management, repository and transaction concepts. Some of these are derived from Open Systems Interconnection (OSI) and Telecommunication Management Networks (TMN).

References

ITU-T. *New Questions assigned to Study Group 10.* COM 10-1-E. World Telecommunication Standardization Conference WTSC Helsinki, 1993.

ITU.T. *Report on the meeting held in Geneva from 19 to 28 October 1993.* COM 10-R 1-E, 1993.

Telektronikk 89(2/3). Kjeller, Norwegian Telecom Research, 1993.

Nørgaard, J et al. PIR 1.3. *Guidelines for service creation and composition.* EU.P103/PIR/1.3. EURESCOM Project EU-P103, 1994.

Colban, E, Christensen, H. *Information modelling concepts.* TB A1.EAC.001 1.0 93. TINA-C, 1993.

Natarajan, N et al. *Information modelling concepts.* TB A.2.NAT.002 1.0 93. TINA-C, 1993.

Singer, N et al. *Engineering modelling concepts.* TB A3.NS.005 1.0 93. TINA-C, 1993.

Message Handling Systems

BY GEIR THORUD

Message Handling Systems (MHS) have been described as "the beginning of the Information Highway". One reason for this is that MHS are the first application oriented data communication service to be implemented on, practically speaking, any type of computer platform, ranging from mainframe computers to mobile hand held devices. A large number of applications have been developed for this infrastructure and an even larger number are under development. The most important ones today are electronic mail, electronic data interchange (EDI) and file transfer. Other emerging applications are store-and-forward fax, voice messaging, work flow automation, electronic forms, computer conferencing and database access.

An important prerequisite for inter-organizational and international electronic messaging is standardization. The most important vendor independent standards are X.400 (also issued as ISO 10021), Internet SMTP (Simple Mail Transfer Protocol) and MHS (which is not the same as MHS described above) in PC LAN environments. This article describes the development of standards in the X.400 area.

Table 1 Organizations and groups working on X.400 or related standards

Organization/group	Area of responsibility
ITU-T Study Group 7 Question 14	Technical standards[1] specifying the Message Transfer Service, all content types and interworking with some other services, e.g. the traditional postal service. The standards are published in the X.400-479 series. Most of them are also published as parts of ISO 10021, as a result of collaboration with ISO/IEC.
ITU-T Study Group 7 Question 18	Protocol Implementation Conformance Specifications (PICS) proforma and standards for conformance testing. The standards are published in the X.480 series.
ITU-T Study Group 7 Question 8	Access to X.400 services via Fax-PAD
ITU-T Study Group 1 Question 12	Standards specifying the service aspects of X.400 based services, e.g. classification of technical functions as mandatory or optional with regard to support by an entity. An example of other service aspects is "recommended transfer times". Standards are published in the F.400 series.
ITU-T Study Group 3 Question 18	Standards specifying principles for accounting applicable to X.400 services and associated applications. The first standard is expected in 1995.
ITU-T Study Group 8 Question 1	Programming Communication Interface (PCI) for Telematic Services providing a common PCI for access to Telefax, Teletex, Telex, and X.400 based services. The X.400 part is currently being developed and will be included in T.611.
ISO/IEC JTC1 SC18 WG4 SWG Messaging	Same area of work as ITU-T Study Group 7 Question 14. The two groups are working as a "collaborative team".
ISO/IEC JTC1 SGFS	International Standardized Profiles (ISPs) specifying levels of support for the most important protocols in ISO 10021 (X.400). The ISPs also include Protocol Implementation Statement proformas. Published as ISO/IEC ISP 10611.
X.400 API Association (XAPIA)	Specifications of Application Programming Interfaces for access to X.400. The most recent specification, "Common Messaging API", may be used for both X.400 and SMTP. It provides a common subset of previous XAPIA specifications, MAPI from Microsoft, and VIM from Lotus.

[1] The correct term for a standard published by ITU-T is "Recommendation".

Brief introduction to X.400

Figure 1 illustrates the fundamental entities in an MHS based on X.400. The Message Transfer Service (MTS) is the backbone of the MHS. It comprises a large number of interconnected Message Transfer Agents (MTA) whose primary function is to provide switching of messages on a store-and-forward basis.

User Agents (UA) give users (persons or computer programs) access to the MTS. There are several types of UAs; each type designed for a specific application of messaging. The most common type is Interpersonal Messaging UAs used for traditional Electronic Mail. In this case, the UA provides, for example, functions for editing, storage of messages and address lists. In addition, it allows the user to request a notification when the message is delivered to the recipient, set transfer priority, etc. Other types of UAs are EDI-UAs and Voice Messaging UAs.

A UA may be directly connected to an MTA, or it may utilize a Message Store (MS) for temporary storage of messages until the UA retrieves them from the MS. Access Units (AU) provide interworking with other services such as Telex, Fax and the traditional Postal service.

All entities mentioned, i.e. MTAs, UAs, MSes, and AUs, are computer programs that may be implemented on almost any type of computer.

When transferred through the MTS a message consists of two primary components, the "envelope" and its "content". The envelope contains parameters, e.g. addresses, which control the processing of the message in MTAs. The content is the information supplied by the user, e.g. text and letter heading. Each type of UA can process one type of content. Content types are today standardized for Interpersonal Messaging, EDI, and Voice Messaging.

Organizations working on X.400 related standards

Work on X.400 and related standards has progressed in several standardization organizations and industry associations. The most important organizations and groups, and their area of responsibility, are listed in Table 1.

Overview of published X.400 standards

Several versions of the X.400 standards have been published since their first introduction by CCITT (now ITU-T) in 1984. The versions are summarized in Table 2. Most standards are published by both ITU-T and ISO/IEC, but a few have been published by only one of the organizations.

In addition to the X-series standards listed in the table, ITU-T has also published a number of F-series standards specifying service aspects and an "X.400 Implementors Guide" which is updated once or twice a year.

Current and future development

The rest of this article will focus on recent developments and future work related to the X.400 series of standards. This work is being pursued by the collaborative team formed by ITU-T SG7 Q14 and ISO/IEC SC18 WG4 SWG Messaging.

Message Store extensions

The capabilities of the MS have been extended to allow storage of submitted and draft messages. A user may utilize several UAs to access the extended MS, e.g. one at work, another one at home and perhaps a third when out travelling. These extensions were approved by ISO/IEC in February 1994, and are expected to be approved by ITU-T in November 1994.

Management of X.400 systems

A number of standards are being developed which specify how the Telecommunication Management network (TMN) may be used to manage entities in an X.400 system. The entities are UAs, MTAs, MSes and Access Units. The management functions offered are Accountability, Security, Configuration, Fault and Performance Management functions. The first standards are scheduled for approval in November 1994, and the others are scheduled for June 1995. Standards describing routing information for use by MTAs are also progressing, but it is not likely that they will be approved before 1996.

Use of X.400 to provide the COMFAX service

The COMFAX service (as specified by ITU-T F.162 and F.163) provides interconnection of facsimile terminals by means of FAX Store-and-Forward units. X.421 specifies how the Interpersonal Messaging Service described in X.420 may be used for interconnection of such units, and is expected to be approved in July 1994.

Access to X.400 through the telephone network

The APS Alliance published the specification of a protocol for provision of the OSI Network Service using traditional modems and the Telephone Network in July 1993. The specification was initially developed for access to X.400, but may also be used to access other OSI based services such as the X.500 based directory. Variants of the specification may also utilize X.28 or X.32 access when the server is connected to a public switched data network This specification is now being progressed as an ITU-T standard, X.445. This is expected to be approved in November 1994.

Table 2 Versions of the X.400 and ISO 10021 standards

Year	Standard	Content
1984	X.400, 401, 408, 411, 420, 430	Specifies the message transfer system (MTS) and the Inter-personal Messaging Service providing an Electronic mail service for transfer of IA5 (ISO 646) encoded text messages and binary files.
1988	X.400, 402, 403, 407, 411, 413, 419, 420	Extends the 1984 functionality with "Distribution lists", the P7 protocol for access to a "Message Store" from PCs, security functions based on cryptography, delivery of messages via the traditional postal service, rules for use of the X.500 Directory Service, and transfer of any (privately) defined type of information.
1990	ISO 10021	Identical to the CCITT 1988 version except for a few very minor extensions.
1991	X.435	Describes a new type of content for transfer of Electronic Data Interchange (EDI) based on, for example, the EDIFACT and ANSI X12 standards. This standard was included in ISO 10021 in 1993.
1992	1988 standards plus X.440. Amendments to ISO 10021.	This is essentially a maintenance version of the 1988/1990 version. The major extensions were a new content type for Voice Messaging (X.440) primarily intended for interconnection of voice-mail systems (accessed via ordinary telephones), enhanced functions for transfer of files, and mechanisms for reverse charging of user initiated receipt notifications.

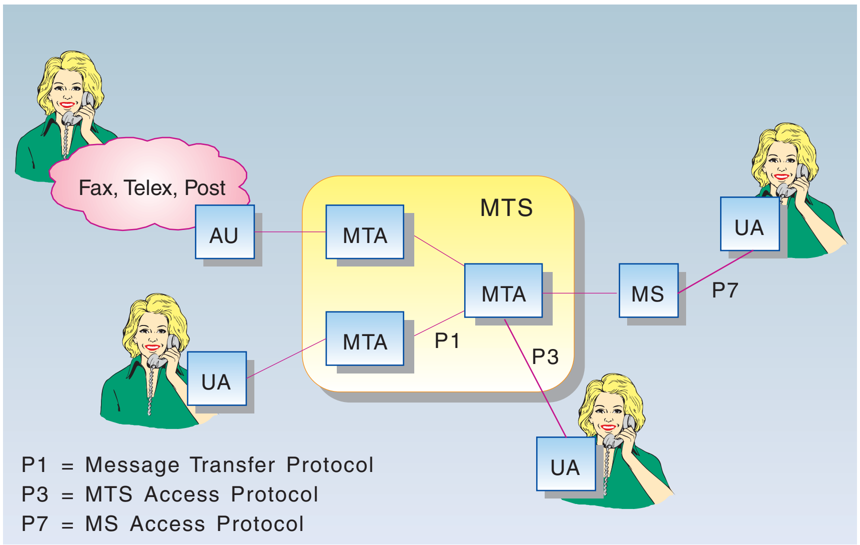


Figure 1 Functional entities in an X.400 based MHS

Extended P3 protocol

The P3 protocol is used for direct interconnection of a UA and an MTA. A new version of the protocol has been approved by ISO/IEC and is expected to be approved by ITU-T in November 1994. The new version provides enhanced security functions and extends the UA's control of functions that are performed by the MTA.

Asynchronous Computer Conferencing (ACC)

ACC is an application of X.400 that allows messages to be exchanged in the context of an "activity" which will often be related to a particular topic or task. An activity has members and may be moderated. Messages related to an activity can be stored on dedicated servers for retrieval at any time. X.400 or special protocols may be used to distribute messages between servers, and by a user to access messages stored on a server. A large number of optional functions have been proposed and it is therefore necessary to specify subsets of the service. One subset is expected to be similar to the USENET News service. So far, the work has progressed slowly. It is unlikely that approved standards will be available before 1997.

Inter-Application Messaging (IAM)

The Society for Worldwide Interbank Financial Telecommunications (SWIFT) has proposed a new content type which is optimized for transfer of files between applications. The content type may carry information describing each file, and for each file an identification of the receiving application. It also provides security functions and mandatory receipt notifications indicating receipt or non-receipt per transferred file. The content type has not received wide support from ITU-T members. If the new content type is approved, it is unlikely to happen before 1996.

Enhancements to X.435 (EDI Messaging)

X.435 may carry EDI interchanges encoded according to the EDIFACT syntax (ISO 9735). A new version of ISO 9735 is expected to be published in 1996, and updates to X.435 are needed for alignment with the new version. In addition, several other enhancements have been proposed. A new version of X.435 can be expected in 1997 if ISO 9735 progresses according to current plans.

Relation between X.400 and Internet standards for messaging

With the recent introduction of commercially offered Internet services, the relation between X.400 and similar protocols used in the Internet has become an important issue. The US member body of ISO/IEC JTC1/SC18 has proposed a "new work item" for preparation of a technical report investigating the possibility of merging X.400 with some of the Internet mail standards. The first formal discussion of this issue will take place in July 1994.

Other extensions

A number of other minor extensions are being developed, but are unlikely to be approved before 1996. Two areas in the work plan are enhancements for multimedia messages and access to X.400 using mobile UAs, but work in these areas has not yet started.

Security

BY SVERRE WALSETH

Introduction

Security is a discipline which may affect many areas of telecommunication. The security functions may be applied to several telecommunication services and on several of the OSI-layers.

The traditional security services includes Authentication, various non-repudiation services, integrity and confidentiality.

- Authentication in the telecommunication context is the verification of the identity of the communication partner.
- Integrity will ensure that any unauthorized alteration of exchanged data is detected by the receiver.
- Non-repudiation of origin resp. receipt means that a particular user, called the originator resp. receiver, cannot repudiate (i.e. deny) to have signed resp. received a particular electronic document.
- Confidentiality is the prevention of unauthorized access to data. This is usually implemented by encryption of the data to be transmitted.

A common feature of almost all security services in an open environment is the fact that realization is only possible through the use of cryptographic techniques. To use these cryptographic techniques some measures must be taken to distribute and update the keys that are needed. For open user groups this problem of key management can be solved by the involvement of a so called Trusted Third Party.

This section contains an overview of international standards regarding the realization of security services related to telecommunications.

The latest results from the following organizations are presented:

- ISO (International Standardization Organization)
- ITU-T (International Telecommunication Union – Telecommunication Standardization Sector)
- UN/EDIFACT (United Nation / Electronic Data Interchange for Administration Commerce and Transport.)
- ETSI (European Telecommunication Standards Institute)
- ECMA (European Computer Manufacturers Association)
- EURESCOM (European Institute for Research and Strategic Studies in Telecommunication)
- NIST (National Institute of Standardization, USA).

ISO/IEC

The main part of Information security work within ISO is done in

- JTC1/SC21 (Information Technology – Information Retrieval, Transfer & Management for OSI)
- JTC1/SC27 (Information Technology – Security Techniques)
- TC68 (Banking Systems).

SC27 is divided into 3 working groups:

- WG1: Requirements, Security Services and Guidelines
  - Identification of application and system requirements
  - Development of standards for security services using techniques and mechanisms developed by WG2
  - Development of supporting documents
- WG2: Security Techniques and Mechanisms
  - Development of techniques and mechanisms which are needed by services identified by WG1.
- WG3: Security Evaluation Criteria
  - Standardization of criteria for security evaluation of IT systems and components.

Ongoing work in WG2:

Entity authentication (ISO/IEC 9798-x)

This standard has 5 parts. Part 1 *(General model)* and part 3 *(Mechanisms using a public key algorithm)* are published as International Standards. Part 2 *(Mechanisms using symmetric encipherment algorithms)* is to be published as an IS this year. Part 4 (“Mechanisms using a cryptographic check function”) is subject to a Draft International Standard (DIS) ballot in 1994. Part 5 (“Mechanisms using zero knowledge techniques”) is expected to reach Committee Draft (CD) level in 1994.

Non-repudiation (Project no. 1.27.06.0x)

The project has 3 parts. Part 2 (“Using symmetric encipherment algorithms”) is being registered as CD, while part 1 (“General model”) and part 3 (“Using asymmetric techniques“) are expected to reach CD level in 1994. All parts are expected to become International Standards no later than 1996.

Digital signatures with appendix (Project no. 1.27.08.0x)

This project has 3 parts; “General model”, “Identity-based mechanisms” and “Certificate-based mechanisms”. They are all at Working Draft (WD) level and are expected to become International Standards by 1997.

Hash-functions (ISO/IEC 10118-x)

Part 1 (“General model”) and part 2 (“Hash-functions using an n-bit block cipher algorithm”) are published as International Standards in 1994. Part 3 (“Dedicated hash-functions”) and part 4 (“Hash-functions using modular arithmetic”) are both WDs and are expected to become International Standards by 1996.

Key management (ISO/IEC 11770-x)

Work on part 1 (“Key management framework”) has just started, and a 1st WD is available. Part 2 (“Mechanisms using symmetric techniques”) and part 3 (“Mechanisms using asymmetric techniques”) are on a CD letter ballot, and they are expected to become International Standards by 1995. Part 4 (“Cryptographic separation”) is a new work item, and a 1st WD is expected in 1994.

ITU-T (earlier CCITT)

Security services may be applied to many of the standards developed by ITU-T. Recommendation X.509 (ISO 9594-8) “Directory System Authentication” is very central for key management issues for various application fields. Other recommendations to be mentioned are:

X.800	(ISO 7498-2) OSI Security Architecture
V4.0	Open Systems Security Infrastructure

UN/EDIFACT

The UN/EDIFACT organization deals with syntax and message development, i.e. only the content of business information to be transferred via telecommunications are candidates for standardization. The UN/EDIFACT organization has a co-operation with ISO, where e.g. the EDIFACT syntax has become an ISO standard (ISO 9735).

The SJWG (Security Joint Working Group) has been assigned the task of recommending standard security procedures to protect EDIFACT messages. The basic pre-requisites of the security solutions are the independence of the transport mechanism and cryptographic algorithm.

The proposed solutions from SJWG includes the AUTACK message, the CIPHER message, message level security and EDIFACT key management.

AUTACK (secure AUThentication and ACKnowledgement) Message.

An AUTACK message used as an authentication message is sent by the originator of messages to facilitate message origin authentication, validation of integrity of content, validation of message sequence integrity or non-repudiation of these messages.

An AUTACK used as an acknowledgement message is sent by the recipient of previously received messages to facilitate confirmation of receipt, validation of integrity of content, validation of completeness or non-repudiation of receipt of these messages.

The AUTACK message can apply to one or more messages from one or more interchanges, or to one or more interchanges.

CIPHER Message

(Cipher is a synonym for cryptology. The message name reflects that the message shall be encrypted or enciphered)

The purpose of the CIPHER message is to provide contents confidentiality for messages and interchanges. The CIPHER message is built by preparing confidentiality security header(s) and by putting the filtered encrypted text inside USE segments.

The CIPHER message is processed by its receivers as any other UN/EDIFACT message. If the receivers are authorized and able to do this, the encrypted USE segments are decrypted and the original entity recovered. This entity should then be processed as intended.

Message Level Security

All services except confidentiality may be provided by the inclusion of generic security header and trailer segment groups after the UNH and before the UNT, in a way which may be applied to any existing message.

Typically, a message security header/trailer pair is required for each security service.

The purpose of the Message Security Header is to specify the security methods applied to the message and to hold the associated data necessary to carry out the validation calculations. A special segment contains details of the security algorithms, other segments may contain the relevant public key certificates.

The Message Security Trailer is used to hold the security result corresponding to the security functions specified in the associated Message Security Header. The Message Security Header and Message Security Trailer are repeated for each set of service and originator. This approach allows for maximum flexibility on future work.

The SJWG’s work on Key Management is performed by a TEDIS project named SAM. The results are presented for SJWG which may give guidelines to the work.

ETSI

Two groups in ETSI are concerned mainly with information security. They are the Security Techniques Advisory Group (STAG) and the Security Algorithms Expert Group (SAGE).

For organizational purposes, STAG appears to be a part of Network Aspects, but STAG is to provide advice and technical assistance to all of ETSI. STAG’s role is to act as the authority and ensure a consistent approach to security standardization throughout ETSI.

SAGE is a closed group within ETSI and its dedicated task is to develop and maintain cryptographic algorithms for ETSI Technical Committees.

In addition, security features are considered in many other ETSI groups. A non-exhaustive list includes:

- GSM. A new export version of the cryptoalgorithm, A5/2, has been developed.
- NA6 Intelligent Network (IN)
- NA7 Universal Personal Telecommunication (UPT)
- SMG5 Universal Mobile Telecommunications System (UMTS)
- TE9 Multi application smart cards. The group has made two specifications, one for a multi application smart card, and one for a security module. A reversible cryptoalgorithm called TESA7 has been developed. This algorithm can be used for authentication, key establishment, key diversification and MAC (Message Authentication Code). In the follow up to of this work, an authentication algorithm to be used with the UPT service has been developed. This algorithm is called USA4.

- TE10 Multimedia and audio-visual services. The project team PT38V has recently specified a cryptographic algorithm to be used in a confidentiality system for audio-visual services.
- RES3 Digital European Cordless Telecommunication (DECT)
- RES6 Trans European Trunked Radio (TETRA).

ECMA

TR-46	Open Systems Security Framework
Std-138	Data Elements and Service Definitions (Security in Open Systems)
WD	Authentication and Security Attribute Service Definition (Security in Open Systems)
WD	Secure Association Service and Management (Security in Open Systems).

EURESCOM

One objective of EURESCOM is to support the development and provision of harmonized pan-European public fixed telecommunications networks and services. Security is an important aspect in this context.

In the work area of telecommunications services, one of the objectives is to provide an overview of the authentication methods applied in telecommunications, analysis of the pros and cons of these, and a guide for developers of services that need authentication.

In the work area of Telecommunications Management Network (TMN), the main objective of the Project P110 “Security and Integrity Requirements”, was to provide a set of recommended security measures for prevention of threats in a pan-European TMN. The final deliverable (Feb. 94) of the project consists of 4 volumes: Main Results, Task Reports, Security Cookbook, and The Crypto Extension.

NIST

NIST has developed a digital signature algorithm called DSA. This algorithm is a standard in the USA for digital signatures (DSS – Digital Signature Standard). NIST has proposed that the algorithm also becomes a European standard. The European countries have adopted a waiting attitude to this question.