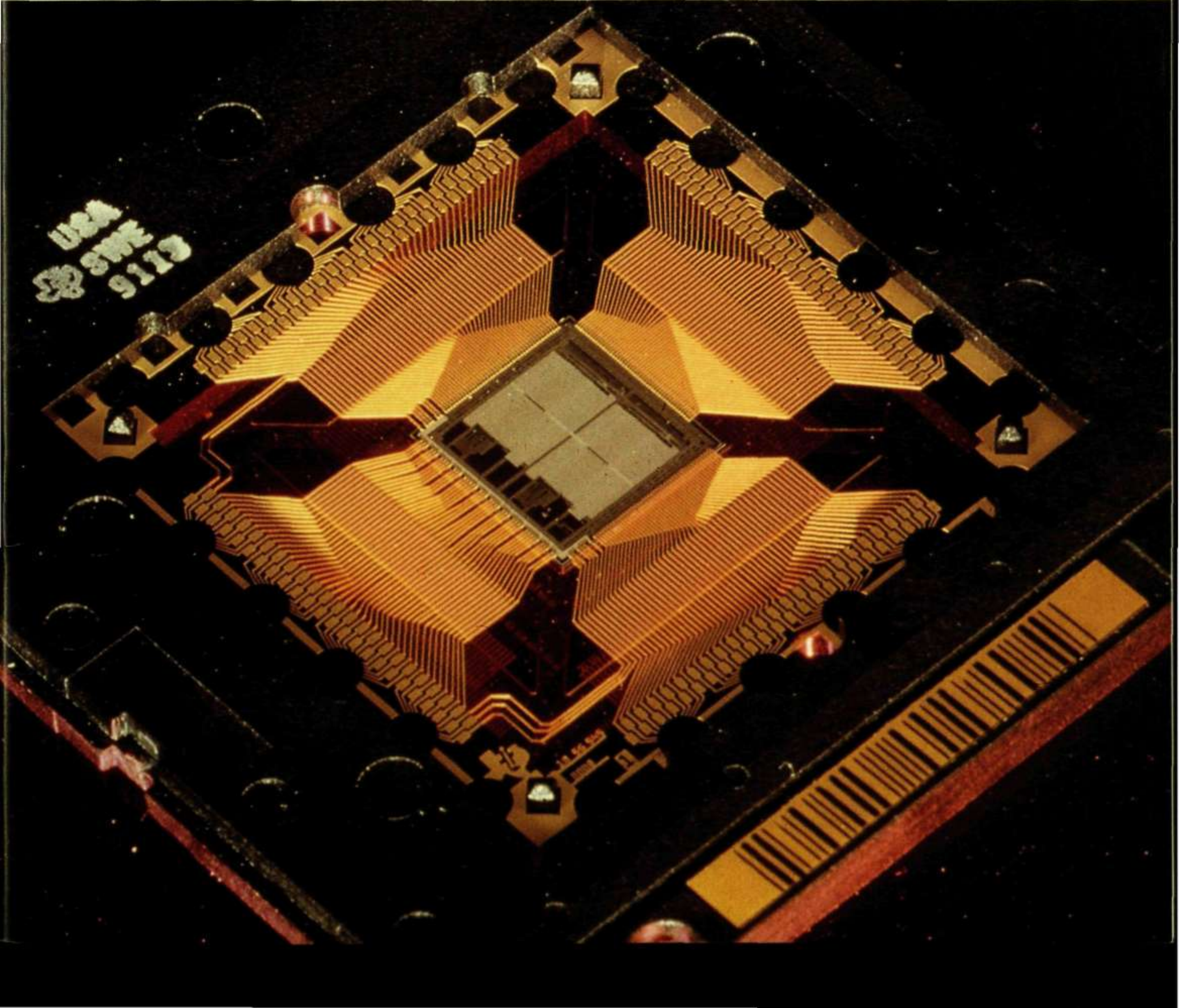
ERICSSON REVIEW

1 1993

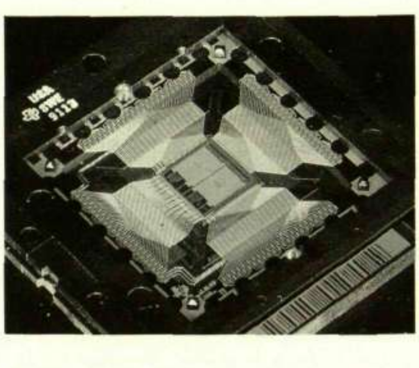


ERICSSON REVIEW

Number 1 · 1993 · Volume 70

Responsible publisher *Bo Hedfors*
Editor *Per-Olof Thyselius*
Editorial staff *Martti Viitaniemi*
Subscription *Eva Karlstein*
Subscription one year $30
Address S-126 25 Stockholm, Sweden
Published in English with four issues per year

Contents



Cover
ATM switch element. One VLSI chip based on sub-micron BICMOS technology has the equivalent of 800,000 transistors in a single TAB package and handles two bidirectional links

The Telecom Evolution in the Broadband Era

Peter Staxén and Claes-Göran Vestin

The emergence of broadband communications coincides with accelerating changes in the climate for provision of telecommunications. The multitude of broadband technologies and the new business climate impose increased requirements on vendors to provide flexible products.
The authors describe several broadband technologies, elaborate on the development of the broadband market and introduce the Ericsson approach to broadband communications.

One of the telecommunications concepts that attracts attention is that of broadband. Several technical concepts are related to broadband, and regardless of which ones will prove successful, it is obvious that broadband communications will play a significant role in our professional and private lives.

Simultaneously, we are witnessing a change in the telecommunications climate. Despite being a public utility, the telecommunications sector is exposed to competitive forces that open doors for new players and their offerings to the market.

Why broadband?

Broadband communications emerged in Local Area Network (LAN) environments with interface electronics and communication protocols supporting data transfer rates of 2, 4, and 8 Mbit/s at an early stage. This development was driven by the advanced applications that were introduced in personal computers and workstations – applications which, in turn, were the result of rapid development in key technologies. Reduced costs of processing power and transmission created the base for distributed processing and information handling.

Public broadband communications became a hot topic when the provision of high-capacity communication channels all the way to the end-user was envisaged. Sitting at his terminal, the end-user would be able to access a whole range of new services including transfer of data, Hi-Fi sound, stills, and moving images. This vision, conceived in the early 80s, was spurred by the fast development of optoelectronics and optical communications.

Together, these technologies facilitated information processing at high speed and the capability to transfer information "in no time".

Thus, the driving forces towards public broadband communications were based on a technical vision, rather than on commercial reality. This fact meant that the development of broadband communications was along different paths and at different paces in the public and private domain.

Fig. 1
Workstations with advanced CAD/CAM applications are heavy consumers of communication bandwidth

Market trends

The assumption that free and fair competition benefits the public more than centralised and planned resource allocation has spurred the deregulation of telecommunications.

The basic trends towards a changed situation on the international telecommunications market can be summarised in the following way:
- The service demand has changed in nature. Customers want mobility (both terminal and personal), bandwidth on demand, access to management functions, flexibility in the establishment of connections, end-to-end connectivity and management.
- It is generally assumed that a competitive market operates more efficiently than a regulated one.

PETER STAXÉN
CLAES-GÖRAN VESTIN
Ericsson Telecom AB

– Equipment and components for the processing and transport of information – particularly digital information – are becoming increasingly cheap.

These trends, and the response from the players involved, have changed the marketplace for both equipment and services, and further changes are to be expected. The two most important consequences are the globalisation of the telecommunications markets and the new competitive environment resulting from deregulation.

The on-going shift – from single-service, single-technology network to multi-technology, multi-service networks – will continue and accelerate. This means that services and technology have to be made even more independent of each other to facilitate reuse of services across different technologies and networks. New business opportunities will be created in the imaginative use of existing networks, but at the same time, gradual upgrading of the infrastructure with broadband capabilities will make it possible to offer a whole range of new services.

As the possibilities of offering customised solutions grow with the technological development, smaller and smaller communities of interest will be identified, and customised solutions will increase in number. This will lead to a fragmentation of the services market, and increasingly diverse requirements will be made on system suppliers and network operators alike.

The introduction of broadband capabilities in various parts of the infrastructure and the offering of broadband services will be step-wise. These capabilities will involve many different technologies, each of them cost-effective in different situations.

Broadband technologies

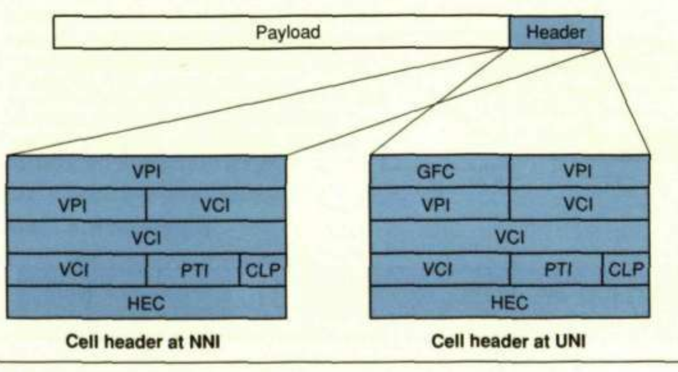
A critical issue related to broadband is the design of efficient and flexible protocols that can support several different services. Two of these protocols, one originally intended for public broadband communications (ATM) and one for private broadband communications (FDDI), are described in the following.

ATM

Originally, Asynchronous Transfer Mode (ATM) was meant to be one of B–ISDN's (Broadband ISDN) cornerstones but, as it turned out, first to be applied in LAN envi-

Fig. 2
The length of the ATM cell is 53 octets. The cell header occupies 5 octets, structured as 6 fields, and contains information required for routing and traffic control. The structure of the cell header is different at the ATM network edge (UNI) and inside the ATM network (NNI). The remaining 48 octets of the cell carry user information

GFC	Generic Flow Control
VPI	Virtual Path Identifier
VCI	Virtual Circuit Identifier
HEC	Header Error Control
PTI	Payload Type identifier
CLP	Cell Loss Priority



Box 1
ATM cell structure

An ATM cell has a fixed length of 53 bytes, or octets, divided into a 5-octet header and a 48-octet body, Fig. 2. The fixed length is a prerequisite for rapid cell processing, which is done in hardware.

As indicated in Fig. 2, the ATM cell header is structured so as to form several fields. The main function of the cell header is to identify the connection and to route the cell from its point of origin to its point of destination. The routing information is contained in the VPI (Virtual Path Identifier) and VCI (Virtual Circuit Identifier) fields. These two fields identify a specific ATM connection: all cells belonging to this connection have the same VPI/VCI values. The header information is protected by a checksum that is contained in the HEC (Header Error Control) field. Since the VPI/VCI values have only local significance, i.e. between two switching points, their value may change as the cell travels through the network and, consequently, the header checksum has to be recalculated at each switching point. The checksum protects the cell against misrouting, but it does not protect the user information carried in the ATM cell. The validation of user data integrity has to be performed on an end-to-end basic by protocols residing on top of the ATM layer.

Box 2
ATM Adaptation Layer (AAL)

Many new services will generate traffic patterns that are suited for immediate transport over ATM. However, ATM is also expected to carry existing services – e.g. voice and frame-based data traffic like X.25 – and for this purpose an ATM Adaptation Layer (AAL) is needed.

The AAL is actually divided into five classes, depending on the characteristics of the traffic that needs to be adapted to the ATM layer. The classification is based on three parameters:
- synchronisation between the sender and receiver (Yes or No)
- information rate (Constant or Variable)
- connection type (Connection-oriented or Connectionless)

Class 1 AAL is intended for circuit emulation, i.e. emulation of constant bit rate (CBR) connection-oriented (CO) circuits for, primarily, voice communication.

Class 2 AAL is applicable to variable bit rate (VBR), CO traffic for video and audio services.

Class 3 AAL is for CO, VBR data traffic like X.25.

Class 4 AAL is for connectionless (CL), VBR data traffic which characterises LAN-to-LAN communication.

Class 5 AAL originates from the ANSI T1S1 committee and is considered to be more efficient for data transfer than Class 3/4 AAL.

In the current set of CCITT recommendations, AAL 3 and 4 are treated as one protocol. A complete set of adopted AAL specifications is contained in CCITT Rec. I.363.

ronment. Its characteristics make it suitable for many applications: backplanes in workstations, LANs, MANs (Metropolitan Area Networks), WANs (Wide Area Networks), multiplexers, concentrators, hubs, switches, etc.

ATM is a cell-based, switched, connection-oriented protocol that combines the good properties of the circuit-switched and packet-switched principles. The choice of ATM, *as a transport mechanism for* B–ISDN, was based on the ATM's capability to provide connections with virtually arbitrary bandwidths and qualities to accommodate a multitude of different services, Box 1. Additionally, ATM offers the possibility of employing statistical multiplexing of connections for efficient utilisation of the transmission resources. ATM was simply considered as the most flexible and robust mechanism for future communications.

Relationship between ATM and STM
The existence of an ATM cell header containing routing information is the main point of distinction between ATM technology and Synchronous Transfer Mode (STM) technology.

In STM systems, based on PDH (Plesiochronous Digital Hierarchy) or SDH (Synchronous Digital Hierarchy) transmission, the available transmission capacity on a link is divided into frames (PDH) or Virtual Containers (SDH). Individual channels are given a specific location – called slot – within such a frame, and information belonging to a channel is carried by this dedicated slot, Fig. 3. Thus, routing is implicitly based on the location of the slot within the frame. Channel bandwidths are allocated in quite large steps.

In contrast, the ATM technology – similarly to other packet oriented transfer modes like X.25 and frame relay – considers link capacity as a medium into which cells can be inserted whenever required. Each cell contains explicit routing information in its header. Cells belonging to different channels contain different routing information in the header and, consequently, they can "travel" along the link in arbitrary order allowing statistical multiplexing of a number of ATM *channels. Only the link capacity*, and procedures for bandwidth allocation, limit the bandwidth that can be allocated to an ATM channel. The utilisation of the link is governed by the number of ATM cells that is inserted into the link per time unit; the larger number, the greater bandwidth is used. These properties of ATM are depicted in Fig. 4.

Resource management and police functions
The inherent ATM flexibility does not come free. Building a network based on ATM requires sophisticated control mechanisms so that the service quality perceived by the network users is maintained.

As in an ordinary telecom network, an ATM connection has to be set up before user information is transferred. Algorithms for connection admittance tend to be complex, and there is still much work to be done in the area of resource allocation in ATM networks.

Once an ATM connection has been set up, and VPI/VCI values have been assigned by the network, it is important to ensure that a network user does not violate the contract established during the set-up procedure. This supervision is required, be-

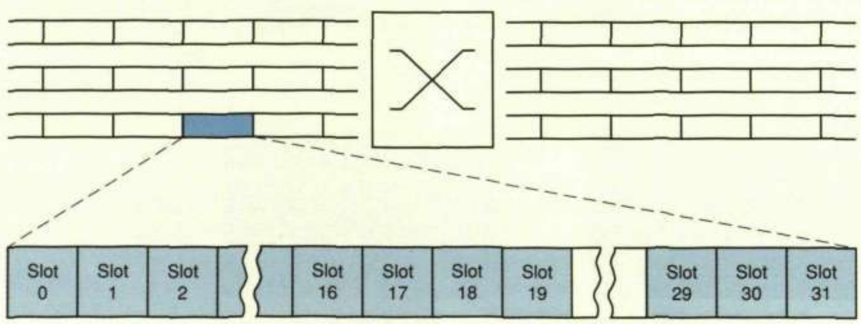


Fig. 3
In STM-oriented transmission and multiplexing schemes, such as PDH and SDH, the available transmission capacity is divided into time slots. Each time slot appears at regular intervals and is allocated to a specific connection. Routing is implicitly based on the location of the time slot within the multiplex structure

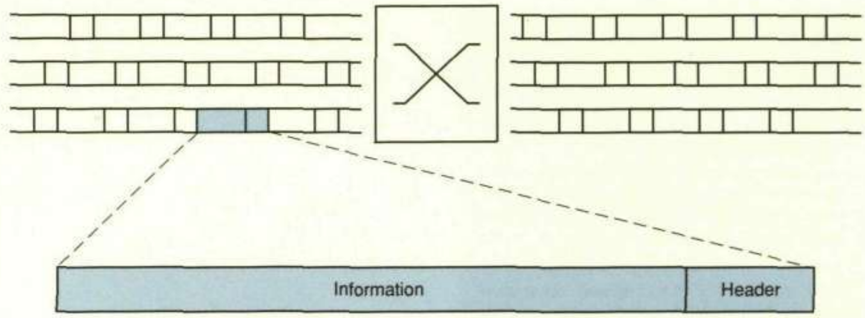


Fig. 4
In labelled multiplexing protocols like ATM, X.25, *and Frame Relay*, the available transmission capacity is shared by several connections. Each information piece is tagged by means of a header. Among several functions, the header identifies the connection. Routing is explicitly based on the information in the header. Cells or packets belonging to the same connection do not appear at regular intervals. This property forms the basis of *statistical multiplexing*

cause a malicious user should not be able to degrade service quality for other users. For this purpose, channel dedicated police functions reside at the ATM network edge, protecting the network and its users, Fig. 5.

FDDI

With the installation of LANs with transmission capacities in the range of 10–16 Mbit/s and the emergence of powerful workstations with high speed interfaces, data communication in large companies and university campuses required a backbone network for interconnection of these items.

A number of different concepts and experimental networks have been conceived over the years. However, only the FDDI (Fibre Distributed Data Interface) and DQDB (Distributed Queue Dual Bus) have attracted commercial interest. The latter one is discussed under the heading of MAN.

FDDI was developed in the US by the ANSI T3X9.5 committee. The purpose of FDDI was to provide a high-speed backbone network for interconnection of LANs, primarily LANs of Ethernet type (IEEE 802.3), in the private domain. Nevertheless, there are public FDDI networks in several countries today, e.g. Finland, Sweden, Germany and the Netherlands.

The FDDI protocol has been expanded to handle isochronous traffic too. This protocol, called FDDI-II, contains functions for supporting 64 kbit/s switched voice. No FDDI-II networks are operational at present, and there are doubts among vendors whether FDDI-II products will ever reach the market, because of the strong push for ATM.

Box 3
FDDI

The name suggests that an optical fibre is used as a transmission medium. Because of the lower costs, multimode fibre was chosen. However, recent advances in coding technology have resulted in standard proposals for FDDI over both Shielded and Unshielded Twisted Pairs. The distance is limited – around 60 meters without repeaters – but it is sufficient for making the terminal attachment to the FDDI backbone significantly cheaper.

The FDDI network consists of dual, counter-rotating rings, each with a transmission capacity of 100 Mbit/s. Normally, only one ring is used for transmission. The other ring is used in faulty situations when the network is reconfigured. The reconfiguration is done automatically by the FDDI management procedures.

Broadband market developments

The market trends towards globalisation and competition have consequences for the provision of broadband communications. The relatively high costs of providing broadband communications, together with competition and scarce capital, will accentuate the differences between the provision of business oriented and mass-market oriented broadband telecommunications.

The provision of broadband communications is fragmented. Different service providers address different segments of the user population. The pace of evolution of broadband communications varies between market segments. These conditions are not expected to change.

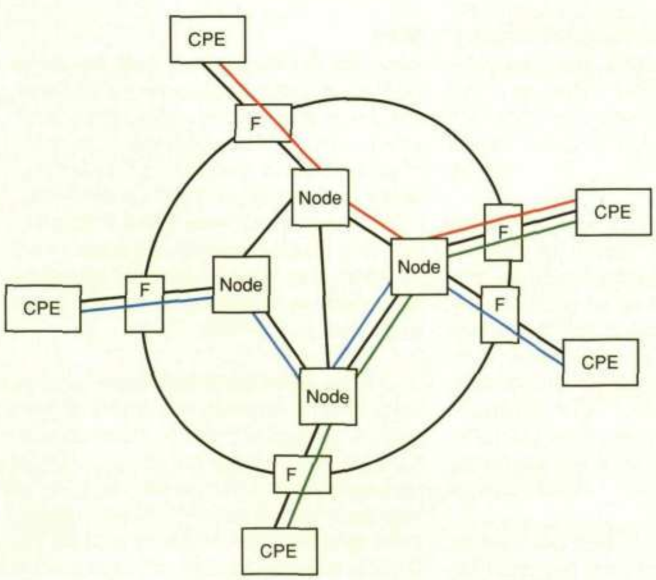


Fig. 5
An ATM network contains a number of interconnected nodes. So-called Virtual Circuits are established between the switching points in the network. An end-to-end connection usually consists of *several concatenated* Virtual Circuits. To protect the ATM network from potential overload situations, police functions are located at the ATM network edge. The police functions ensure that no connection violates the conditions agreed when connections are set up

CPE Customer Premisses Equipment
Node ATM Switching node
F Flow Enforcement (Police function)
Virtual connection 1
Virtual connection 2
Virtual connection 3

Box 4
ATM Forum:

The primary objectives of the ATM Forum are:
- definition of the technology through the technical committee's activities
- education of the industry through the publication of articles and tutorials, advertising, creation of a speakers bureau, promotion of end-user groups and other activities.

So far, the ATM Forum has agreed on the specifications of the ATM User-Network Interface (UNI), which covers the permanent, i.e. non-switched, Virtual Path/Circuit service over both private and public UNI, including:
- initial service attributes over the two UNIs. (Public UNI refers to the interface at the TB reference point, while Private UNI refers to the interfaces at either RB or SB reference points (see B–ISDN below).
- four physical layer specifications intended for the carriage of ATM cells, namely DS3, SONET STS-3c, FDDI and 155 Mbit/s Multimode Fibre Interface. (Other physical layers to be specified include SONET STS-3c over Shielded and Unshielded Twisted Pair and the PDH hierarchy excluding the lowest rates).
- ATM layer specification common to all physical layers (identical with CCITT's specifications)
- Interim Local Management Interface (ILMI)

In addition, ATM Forum is planning to specify the UNI signalling protocols, traffic management procedures and inter-carrier interface. Other interests concern interworking issues between ATM and Frame Relay and SMDS respectively. There are also plans for establishing a European branch of ATM Forum.

The ATM Forum
A significant catalyst for promoting broadband communications in general and ATM in particular is the ATM Forum. It was formed in the US as an international consortium chartered to accelerate the use of products based on ATM technology. The ATM Forum is not a standards body but it informs standardisation bodies like ANSI and CCITT about the progress of its work, Box 4, through contributions made by individual member organisations.

The member list comprises over 160 organisations, including Ericsson, representing various industry sectors: LAN and WAN vendors, vendors of interworking products, computer vendors, CPE (Customer Premises Equipment) vendors, switch vendors, public telecom operators, other carriers, semiconductor manufacturers, government agencies, research and user organisations. The very active participation of CPE vendors is significant and is a prerequisite for the success of ATM in the business environment.

Business-oriented broadband communications
The immediate need of the business community is to transfer large amount of data between computers and workstations at different sites. Several network concepts and services have been introduced to meet this need, which has arisen due to the expanding use of the client-server type of data applications.

MAN
The idea of Metropolitan Area Networks (MANs) is a natural consequence of the increased deployment of LANs among large private and public institutions, the growth of geographically distributed sites belonging to the same organisation, and the desire to interconnect these LANs in a seamless way, Box 5. However, the original idea of MANs also incorporated the capability to provide isochronous traffic, like voice and video, over MANs.

In a sense, MAN and B-ISDN address similar problems although with a different approach, stemming from the different traditions of the bodies that develop these concepts. Apart from the difference in the approach to the problem there are also more tangible differences. First of all, the B-ISDN concept defines interfaces, while the MAN concept specifies a network architecture. Furthermore, functions like operation, administration and maintenance, charging and interworking are not of prime concern in the MAN world although the concept itself is applicable both in the private and in the public domain.

SMDS
The MAN concept is fundamental to the Switched Multi-megabit Data Service, SMDS, defined by Bellcore, which addresses the high-end of the market for high-speed data communications. SMDS is a public, switched, connectionless data service aimed at providing LAN-like performance over a wide area. Several access classes are offered, ranging between 1.5 and 45 Mbit/s.

The SMDS specifications are based on the DQDB protocol but, basically, the service may also be provided using ATM as the transport mechanism. In fact, some of the SMDS field trials being conducted use ATM based switches.

A European version of SMDS called CBDS (Connectionless Broadband Data Service) is standardised by ETSI. However, the success of SMDS/CBDS will depend very much on the costs associated with the service and the alternative means of satisfying the user demand for high speed datacom.

Frame Relay
While the SMDS addresses the high-end of the broadband data communications market, the Frame Relay service addresses the low-end. Frame Relay is actually derived from ISDN where it is referred to as the Frame Mode Bearer Service (FMBS, CCITT Rec. I.122). It received intensified attention when it was applied to LAN interconnection. Frame Relay utilises X.25-type frames with two additional octets containing addressing information and Cyclic Redundancy Check information.

Frame Relay has a throughput which is significantly higher than traditional X.25 networks, because relaying is performed at frame level instead of packet level. However, current specifications limit the maximum speed to 2 Mbit/s, which is insufficient for many applications. Like ATM, only semi-permanent connections are supported by current standards, which limits the flexibility of Frame Relay. Commercial

Box 5.
MAN

The MAN initiative was conceived in the IEEE 802 committee on standards for LANs. The purpose was to define a protocol that provided transmission capacities in excess of 100 Mbit/s and could be used over large geographical areas, typically 200 km in diameter. This protocol should also conform to other LAN 802.x family of protocols (802.3 CSMA/CD, 802.4 Token Bus, 802.5 Token Ring) as to the services it provided.

The protocol chosen for the MANs, designated 802.6 or DQDB (Distributed Queue Dual Bus), implements a distributed queue, maintained in each node, which is used for determining which node has access to the transmission medium. The medium itself consists of two, counter-directed, unidirectional, slotted buses. The DQDB slot has the same size as the ATM cell but the slot header is slightly different from that of ATM. Nevertheless, interworking between DQDB-based subnetworks and ATM networks is considered to be a minor problem.

Frame Relay offerings today can be found mainly in the US.

Multimedia

Apart from data communications, there is also a driving force for broadband communication referred to as "multimedia". Multimedia implies that several information types are simultaneously accessible to the user who can control their presence and the information contents in an interactive way. With this definition, TV is not considered a multimedia service since the user cannot control the contents; except by switching to another channel, of course.

The multimedia technology has a wide scope of applications such as education, remote purchasing of goods and real-estate, newspapers, tourist guides, etc.

An important aspect of multimedia services is the terminal itself. It is expected that a personal computer equipped with appropriate hardware and software will be the first kind of multimedia terminal. Already we can see product announcements which are prepared to support multimedia. However, these announcement only refer to local capabilities, i.e. attachment of special units directly to the personal computer. Provision of public-networked multimedia services is still several years ahead. On the other hand, LAN-network multimedia may be just around the corner.

Mass-market broadband communications

The planning of mass-market public broadband communications began at the time when monopoly related to provision of networks and services was still considered a natural state of affairs and resulted in concepts like B–ISDN. Although this concept may never be realised in the way it was intended, it has had a tremendous impact on the development of broadband communications.

The driving forces for mass-market broadband communications are access to information and entertainment. Access to information represents a unification of telecom, datacom and media. Our society is increasingly dependent not only on access to information but also on intelligent screening of the vast amount of information produced every minute. It is in this domain that an explosion of services is foreseen. The media industry will produce information, the datacom industry will provide the terminals and mechanisms for selecting the information, and the telecom industry will provide the means of conveying the information to/from users.

On the entertainment side, TV distribution, video-on-demand type of services, interactive networked multi-user games etc might prove popular if attractively priced. It is interesting to note that although several years ago distribution of High Definition Television (HDTV) was the service that required most bandwidth, this is no longer the case due to improved coding techniques.

B–ISDN

Broadband ISDN, or B–ISDN for short, standardised by the CCITT, was originally considered as an extension of the narrowband ISDN. In reality, the only features the two concepts have in common are the reference model containing the functional groupings NT1, NT2, TE and TA and the reference points between these functional groupings. Fig. 6 illustrates the B–ISDN reference model.

Instead, B–ISDN is an infrastructural concept implying deployment of optical fibres

Fig. 6
The B–ISDN reference model is identical with the ISDN reference model. It contains the same functional groupings and reference points. The B–ISDN UNI runs across the T_B reference point. In some configurations, the S_B and T_B reference points may coincide. The B–ISDN NNI is a transmission interface inside the B–ISDN network

B-NT1	Broadband Network termination 1
B-NT2	Broadband Network termination 2
B-TA	Broadband Terminal Adapter
B-TE1	Broadband Terminal Equipment (Type 1)
B-TE2	Broadband Terminal Equipment (Type 2)
NN	Network Node
UNI	User Network Interface
NNI	Network Node Interface
S_B, T_B, R_B	ISDN Interfaces

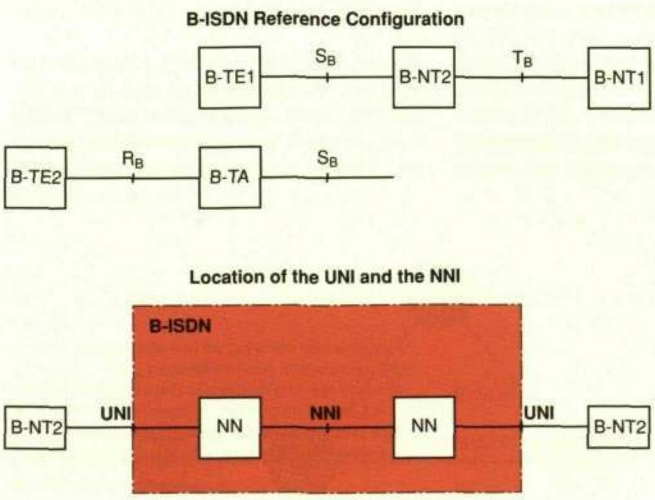
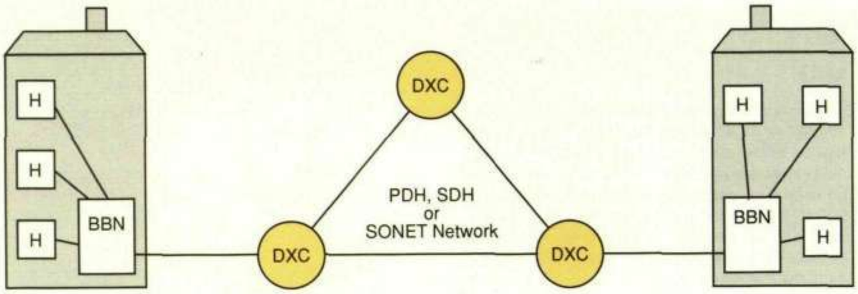


Fig. 7.
At the initial stages of ATM-based networks, *leased lines based either on PDH or SDH will* interconnect ATM islands. These islands will be located on the premisses of the users. Already at this stage, both users and operators will benefit from the capabilities offered by the network

H	Floor Hub
DXC	Digital Cross Connect
BBN	Broadband Business Node



all the way to the customers' location, new switches, new signalling and transport protocols and new services which will take advantage of this new infrastructure.

One of the most important facets of B–ISDN is the specification of the two interfaces: UNI and NNI.

The UNI (User-Network Interface) is the interface through which users attach their terminals to the B–ISDN. In this context, the terminal should be considered in a rather broad sense of the word, ranging in complexity from simple telephones over PABXs to LANs. The NNI (Network Node Interface) is the interface between two network nodes. The NNI specifications agree with the CCITT Rec. G.707-G.709, which specify SDH rates, formats and multiplexing structures.

Introduction of Broadband Capabilities

As indicated above, the deployment of broadband will be demand-driven and the scenarios for describing broadband introduction should therefore be seen as a reflection of the current state of matters. With this in mind, the following scenarios can be drawn up:

Phase 1: Business ATM over PDH/SDH network
Broadband capabilities based on ATM will first be introduced in business networks. ATM- based LANs and hubs are already operational. These ATM islands will be interconnected via dedicated links in the existing networks, using either PDH or SDH as the underlying transmission method, Fig. 7.

This capability provides some significant advantages to both corporate users and network operators. Corporate users will gain access to high-speed wide-area communications, while network operators will gain from the more efficient and flexible use of the transport network resources if SDH (SONET) is used.

Phase 2: ATM enhancement of the infrastructure
As the traffic volume increases, due to a larger number of ATM islands, the broadband infrastructure initiated in phase 1 is extended with ATM cross-connect systems that will increase the flexibility of the network, Fig. 8.

Although the primary role of the ATM transport network is to carry broadband traffic, it can also be used as an efficient medium for narrowband traffic using circuit emulation. Virtual Leased Lines for corporate users will be provided through 1.5/2 and 34/45 Mbit/s circuit emulation functions in the ATM nodes. Customer control over the virtual resources will be provided via management interfaces. Since the tariffs will probably be based on throughput rather than capacity, there is a potential for lower communication costs.

Phase 3: Connectionless broadband data communications
Once the ATM transport nodes and transmission 'pipes' are in place, and the demand for switched data services increases, the transport nodes can be equipped with connectionless servers for SMDS/CBDS to provide switched data communications between business premises, Fig. 9.

Any of the ATM nodes with cross-connect functionality can be upgraded to a multiservice node, catering for both Virtual Leased Lines and connectionless services. The emerging demand will dictate

Fig. 8.
Further flexibility in operation and utilisation of the ATM network is achieved by introducing ATM cross-connect functionality. It will also be possible to handle *narrowband traffic via so-called* Circuit Emulation functions in the ATM network nodes

C	CPE
AN	ATM Access Node
NN	ATM Network Node

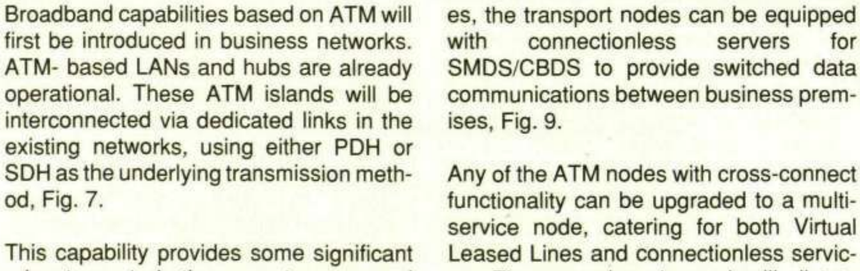
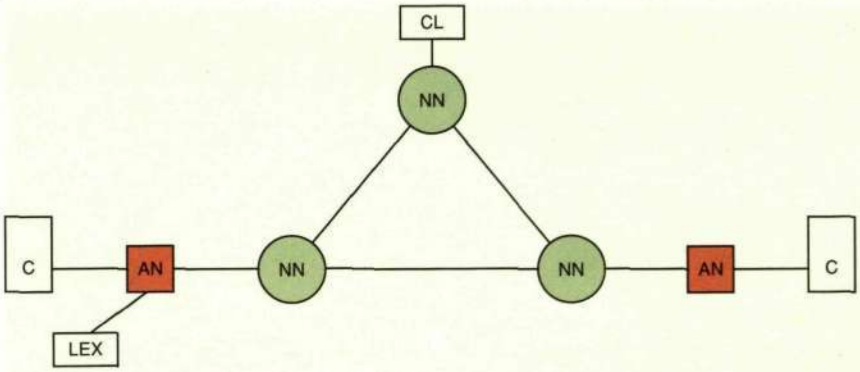


Fig. 9.
By introducing connectionless servers, attached to the ATM cross-connect nodes, switched ATM traffic will be handled. This kind of traffic will appear either as out-sourced or overflow traffic from private networks. The servers need only be introduced in those ATM network nodes where it is economically motivated

CL Connectionless server
LEX Local Exchange
C CPE



which nodes will be upgraded and at what pace.

Phase 4: Introduction of signalling and routing
The functionality of the ATM network will be further enhanced by introducing signalling and routing capabilities, Fig. 10. A connection oriented signalling link allows per-call switching of virtual channels in ATM.

Users can initiate and terminate calls as needed, using whatever bandwidth is required, within the terms of agreement with the network.

Phase 5: Broadband communication to the residential market
The previous phases were aimed at the business communications.The roll-out of broadband services to residential users will emerge later, due to the necessary reconstruction of the access network. To justify the associated costs, new attractive services have to be offered to the residential users.

The Ericsson approach to broadband communications

The primary objective for Ericsson's broadband development program is a timely and cost-effective introduction of broadband capabilities, helping the network and service providers to respond to market demands in a timely fashion.

With these objectives in mind, Ericsson has developed two concepts that form the pillars of the broadband program: the ATM Pipe Switch and the Generic Broadband Module (GBM). Together, these two offer a unique flexibility that will considerably increase the robustness of the network and reduce network management costs. They also facilitate the ability to upgrade and reconfigure the network elements easily, to meet changing demands. A description of the ATM Pipe Switch is given in a separate article in this paper[1].

The Generic Broadband Module (GBM)

The broadband network architecture, as envisaged by Ericsson, is based on several key building blocks. These building blocks are all based on the same technological platform: the Generic Broadband Module.

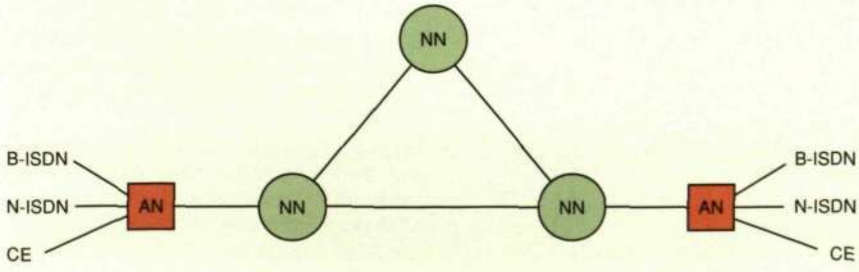
Different GBMs provide different functions, but the basic configuration includes an ATM switch core, a processor board and device boards, such as access boards, server boards or signalling termination boards. The principal structure of the GBM is depicted in Fig. 11.

All the printed board assemblies are packaged in a GBM subrack. By changing or adding boards and software packages, the broadband functionality of the GBM can be varied and evolve in line with progress in technology and changing market requirements.

The GBM uses a well-defined internal structure and interfaces to create a modular system with flexible use of hardware units. GBM-based products with different switching rates can be interconnected directly. This will also be the case with different generations of switch equipment. All

Fig. 10.
User-initiated calls will be handled when signalling and routing capabilities are introduced. Signalling protocols will allow users to set up calls with various connection topologies (point-to-point, point-to-multipoint), bandwidths and other parameters associated with quality of service

CE PDH Circuit Emulator



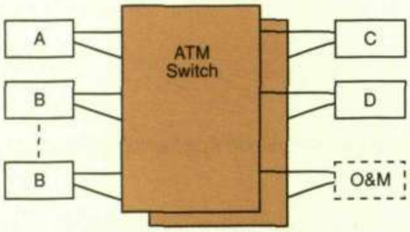


Fig. 11.
The Generic Broadband Module (GBM) is a modular platform for building network nodes with varying broadband capabilities. A GBM can be equipped with several types of printed board assemblies: ATM switch core, processor boards, access boards, server boards and signalling termination boards

A	Signalling Termination board
B	Access board
C	Server board
D	Processor board
O&M	Operation and Maintenance Support

boards in a GBM are self-identifying and can be plugged into any available slot. The system is self-configuring, which makes it easy to handle and improves its robustness.

A single GBM can be used as a small broadband node, or be combined with other GBMs to implement larger broadband nodes of almost any size. Notably, the node concept is distributed, and a single node can consist of several GMBs spread over a wide geographical area.

The first implementation of the GBM offers 155 Mbit/s access to the ATM switch core. The core itself can be configured as a multiplexer, a concentrator or a switch.

When high switching capacity is needed, a unit with only ATM switch boards is used, physically separated from the access devices, etc. which are housed in interconnected GBMs.

Ericsson Broadband Products

Ericsson's ongoing design of broadband products reflects the company's strategy for introducing broadband capabilities into the network. This strategy is based on three concepts:
- responsiveness to market demand
- adherence to standards
- reusage of existing infrastructure.

Based on the GBM and ATM Pipe Switch, several products aimed for both public and private networks are under development.

For public networks, various types of transport, switching and access nodes can be configured. Transport nodes encompass ATM cross-connects, ATM concentrators and ATM multiplexers.

Transition to an ATM-based network is facilitated by the inclusion of circuit emulation functions in the Ericsson ATM node concept. With these functions, transmission circuits like G.703 can be emulated by the ATM network. Management of the ATM transport network will be performed by the FMAS (Facility Management System) which is based on the TMOS platform. Management of ATM transport network elements is performed via object-oriented interfaces following the TMN principles.

Switching nodes can terminate higher order protocols, such as signalling or data protocols. The first application will be a connectionless server for supporting high-performance LAN interconnection. The user of this service will be supported by advanced customer control functions, allowing access to fault, performance, configuration, accounting and security management, all provided via TMOS.

Access nodes will provide various interfaces to both broadband and narrowband networks. Three types of interfaces to customer environment are envisaged:
- Circuit emulation access for narrowband and wideband traffic into ATM
- ATM access
- Frame Relay protocol emulation access.

Broadband access nodes can be interconnected over wide-area transport networks using either STM or ATM technology.

For private networks, several types of nodes are designed, such as the Broadband Business Node (BBN) and an ATM hub. A more detailed description of different solutions for both public and private networks will be presented in a coming issue of the Ericsson Review.

Conclusions

The emergence of broadband communications coincides with the opening of the telecom market to new players. This new market climate, together with a multitude of possible broadband technologies, requires a new approach to address the broadband market.

From being a technology originally aimed for public broadband networks, ATM has

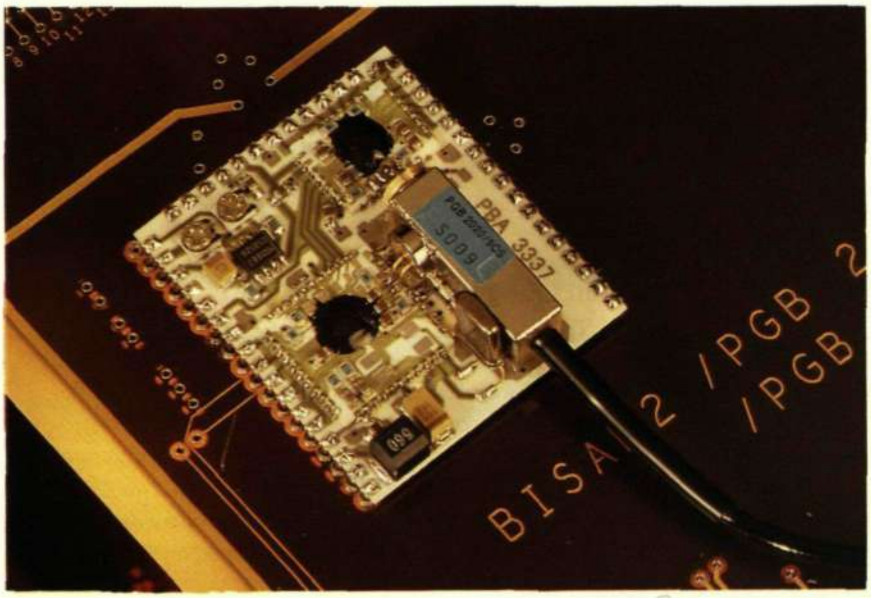


Fig. 12
Broadband communication is to a large extent based on use of optical fibres. The picture shows *an opto-link connected to a circuit board*

first emerged in private LAN environment. This development is driven by requirements for high-speed data communications in both LAN and WAN environments. The flexibility of the ATM technology makes it ideally suited for these applications.

The introduction of broadband capabilities into networks will be evolutionary, following the demand from both the business community and the public. An overlay broadband network will interwork with the existing networks.

The broadband strategy adopted by Ericsson aims at gradual introduction of broadband capabilities, both in terms of capacity and functionality. With a modular, flexible and open broadband system structure, based on the GBM and ATM Pipe Switch concepts, Ericsson intends to provide flexible, cost-efficient and complete network solutions to the benefit of all users.

References

1 Larsson M., Ljungberg M. and Rooth J.: *The ATM Switch Concept and the ATM Pipe Switch.* Ericsson Review *70* (1993):1, pp. 12–20.
2 Danielsson S.:*An Introduction to the Ericsson Transport Network Architecture.* Ericsson Review *69* (1992):3, pp. 58-77.
3 Tarle, H.: *FMAS - An Operations Support System for Transport Networks.* Ericsson Review *67* (1990):4, pp. 163-182.

CCITT Recommendations
I.150 B–ISDN Asynchronous Transfer Mode Functional Characteristics
I.361 B–ISDN ATM Layer Specification
I.363 B–ISDN ATM Adaptation Layer (AAL) Specification
I.371 Traffic Control and Congestion Control in B-ISDN
I.413 B–ISDN User-Network Interface
I.432 B–ISDN User-Network Interface Physical Layer Specification
G.707 Synchronous Digital Hierarchy Bit Rates
G.708 Network Node Interface for the Synchronous Digital Hierarchy
G.709 Synchronous Multiplexing Structure

Bellcore Specifications
TR-TSY-000772 Generic System Requirements in Support of SMDS Service
TR-TSY-000773 Local Access System Generic Requirements, Objectives and Interfaces in Support of SMDS Service

ANSI Recommendations
T1S1/92-283 Proposed Protocol for AAL-5 Common Part
X3.139 Fibre Distributed Data Interface (FDDI) Token Ring Media Access Control (MAC)
X3.166 Fibre Distributed Data Interface (FDDI) Physical Layer Medium Dependent (PMD)
IEEE 802.6 Distributed Queue Dual Bus (MAN)

ATM Forum Recommendations
ATM User-Network Interface Specification, Version 2.0,

The ATM Switch Concept and the ATM Pipe Switch

Mikael Larsson, Martin Ljungberg and Jan Rooth

Asynchronous Transfer Mode, ATM, will be an essential bearer service in the next generation of broadband telecommunication networks. A major reason for this is ATM's superior capacity for offering flexible bandwidth, as compared with the multiplexing techniques used in today's Plesiochronous and Synchronous hierarchies. There is an emerging demand for services to support users with various needs of bandwidth, across a common user interface and a common and simple bearer service through the network. To meet this demand, Ericsson has developed a framework for switches based on ATM technology.

The authors describe the principles of the ATM Switch concept and a specific implementation of an ATM switch, called the ATM Pipe Switch.

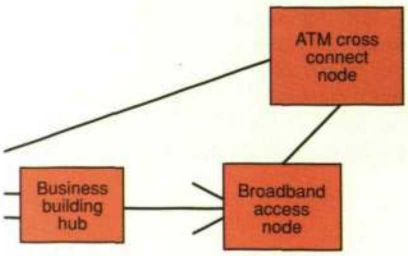


Fig. 1
The ATM switch can be used in different network configurations –.as a building hub on customers' premises or as a public cross- connect, supporting services with widely differing requirements

Expected ATM services
VLL Virtual Leased Lines
FR Frame Relay
SMDS Switched Multimegabit Data Service
CE Circuit Emulation

A key building block in networks based on ATM technology is the ATM switch. The switch will be used in different environments for private and public telecommunications and must be capable of supporting different applications and different network configurations. This means that the switch must be flexible with respect to size, bandwidth, reliability and performance. Fig. 1 illustrates some ATM switch applications.

Ericsson has developed a general-purpose ATM switch. Thanks to its modular design it can be configured to work in different locations in a network and to support a number of broadband applications. Two important guidelines have been followed in the design of the ATM switch:

- Separation of the switching function from the broadband applications by means of well defined interfaces, in order to be able to enhance the switching capability independently of the application
- Use of a well defined interface between the hardware-dependent and hardware-independent software parts of the switching system, to allow the software to be easily upgraded and the hardware to be developed in pace with technological advances.

These guidelines were first applied to a general architecture for the ATM switch, called the ATM Switch concept. In accordance with this concept, and drawing on the experience gained by two generations of ATM system and technology prototyping, a hardware design – the ATM Pipe Switch – has been defined.

The ATM Pipe Switch is based on a quadratic switch core which is connected to a number of switch ports. The switch core is non-blocking and designed to handle virtual connections with various bit rates. The switch core can handle both point-to-point and point-to-multipoint connections. The Pipe Switch supports both cell-delay and

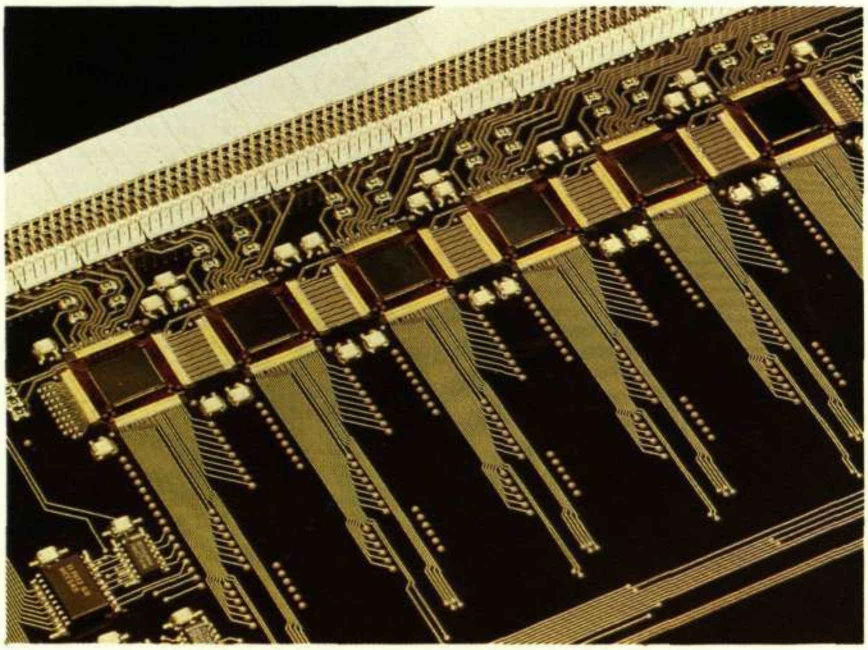
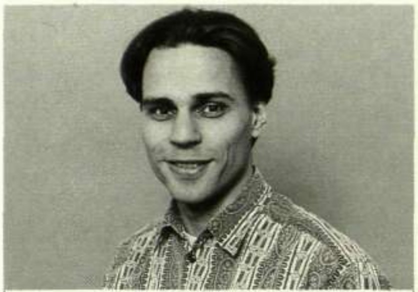


Fig. 2
The ATM switch matrix

MIKAEL LARSSON
MARTIN LJUNGBERG
JAN ROOTH
Ellemtel Utvecklings Aktiebolag

Box 1

Priorities and Quality of Service (QOS)

In the standard ATM layer[1] there is provision for a cell-loss priority mechanism on a cell-to-cell basis. Each cell can be marked with one of two priority levels. Cells marked with high-loss priority should be less likely to be discarded, due to buffer overfill, than cells marked with low-loss priority. A threshold function is therefore used in the Pipe Switch buffers. When the buffer contents are above this threshold, only cells with high cell-loss priority are accepted and buffered. All the others are discarded.

Connections in the ATM layer can be set up with different QOS (Quality of Service) classes. A QOS class is characterised by parameters for cell-loss probability and maximum cell delay. To support different QOS classes, the Pipe Switch provides both cell-delay and cell-loss priority on a per connection basis.

Delay priorities are achieved through two parallel queues in the egress buffers. Cells that belong to high-priority cell-delay connections are placed in one queue and all other cells in the other. The high-priority queue is always served first, and cells from the other queue are sent to the output only when the high-priority queue is empty. This means that the delay for cells belonging to high-priority connections are independent of the number of cells waiting for retransmission in the low-priority buffer.

The Pipe Switch also supports cell-loss priorities on a connection-to-connection basis. These priorities and those on a cell-to-cell basis are handled in a similar manner. There are three levels of connection-oriented cell-loss priorities, with one threshold for each level. When the limit in the buffer for cells belonging to a certain priority class is reached, incoming cells of this category is discarded as long as the buffer-overfill condition prevails.

Another means of traffic control provided by the ATM layer is the Explicit Forward Congestion Indication, EFCI, carried by a bit in the ATM header. In the Pipe Switch, this bit is set in outgoing cells if the buffer level is above a certain threshold, different for the different classes of service. This mechanism indicates to the terminating side of a connection whether or not there is congestion - or close to congestion - anywhere along the route.

cell-loss priorities, Box 1, and thus complies with different QOS (Quality Of Service) specifications for different services.

The design of the quadratic switch core is based on a pipe structure, permitting a flexible and modular design of switches to meet different capacity demands.

Two full custom circuits designed with 0.8 μm BiCMOS technology – the ATM Input Circuit, AIC, and the ATM Output Circuit, AOC – form the platform for a large group of switch products, ranging from small switches and concentrators with 20 connections for 155 Mbit/s, up to switches with 80 Gbit/s capacity, Box 2. The access ports accommodate speeds from 155 Mbit/s to 2.5 Gbit/s.

The ATM switch can be configured to meet the requirements of a small business hub as well as a large cross-connect in the public network. Different services with various capacities, quality and bandwidth requirements can be provided by the system. The ATM Switch concept is structured in such a way that new versions of software and hardware can be added to the system in a flexible and modular fashion.

The ATM Switch Concept

Hardware architecture

The hardware architecture of the ATM switch defines two main building blocks: the Switch Core, SC, which performs space switching, and the Switch Port, SP, which represent the inlet and outlet functions. SC forms a separate unit, whereas SP is located on the same printed board assembly as the different access devices that handle the applications and the Processor Station, PS, where the control functions are performed.

To obtain a system which permits the different entities to be upgraded indepen-

Box 2

Facts about AIC and AOC.

AIC and AOC are two full custom ASICs developed for use in the Pipe Switch. Both circuits are produced in a 0.8 micrometer BiCMOS process.

AIC is the smaller of the two circuits. It has a total transistor count of around 0.5 million and an area of 140 mm^2. Specially designed bit synchronisers are used for phase alignment of the incoming data streams.

AOC, the larger circuit, has a transistor count of 2.2 million. A major part consists of a 310 Kbit embedded RAM, which is used for cell buffering. The area of the AOC is approximately 190 mm^2.

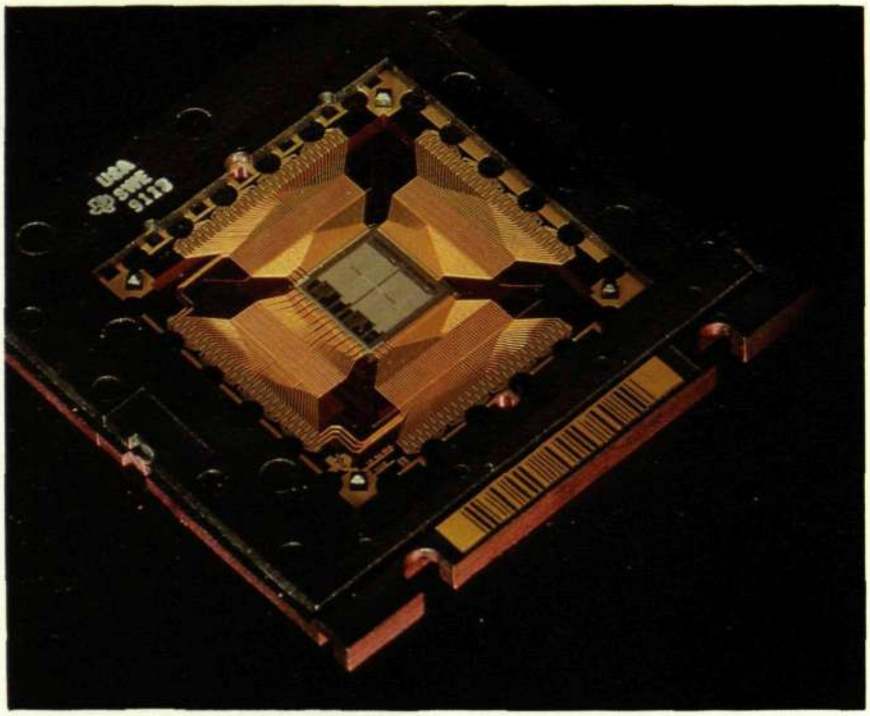


Fig. 3
A single TAB package, containing the equivalent of 800,000 transistors in sub-micron BIMOS technology, handles two bidirectional links

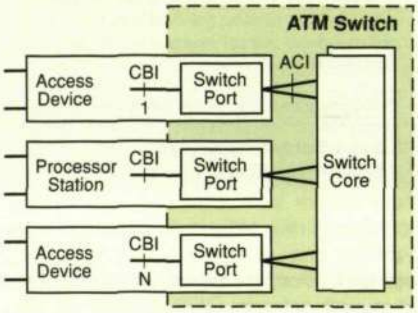


Fig. 4
The ATM Switch Architecture is based on two main building blocks: Switch Core and Switch Port. The integrated Switch Ports are located on the same printed board assemblies as the access devices. The CBI (Cell Bearer Interface) is a general interface for all types of application-dependent hardware. The ACI (ATM Core Interface) is an internal interface between the Switch Core and the Switch Ports

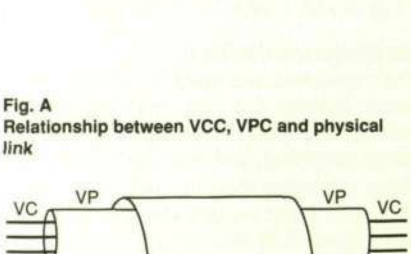
dently, the ATM switch is encapsulated by well defined hardware and software interfaces. This means that the functions performed in the ATM switch are strictly separated from those performed by other node devices.

Fig. 4 depicts the ATM switch and its physical interfaces. In the figure, two access devices and one Processor Station (PS) are shown – each connected to the Switch Core via its own Switch Port. The Cell Bearer Interface, CBI, connects the access devices to the Switch Ports. The Switch Core and the Switch Port are linked by the ATM Core Interface, ACI.

Switch Ports

The Switch Ports form an adaptation between the access devices and the Switch

Fig. A
Relationship between VCC, VPC and physical link



Box 3

Virtual Path and Virtual Channel Connections (VPC and VCC)

ATM is a connection-oriented packet service, where the information is carried by ATM cells. Two types of connections are used: Virtual Path Connections (VPC) and Virtual Channel Connections (VCC). Cells belonging to a VPC are identified by the VPI (Virtual Path Identifier) in the ATM cell header. The connection is regarded as virtual, since it shares its physical media with other connections. A physical link may contain 4096 VPCs, at most, and the User Network Interface no more than 256 VPCs. The VPI is unique only for a specific physical link.

Just as physical links may carry a varying number of Virtual Paths, a Virtual Path may carry up to 64 k Virtual Channels. A VCC is identified by the combined information of the VPI and VCI fields in the ATM cell header. The VCI is unique only for a specific VPC, but the VPI together with the VCI are unique for a physical link, Fig. A.

The VPCs and the VCCs form two different network layers: the VPCs are related to the lower layer and the VCCs to the higher one. In the case of VP switching, the VCI field is considered as user-data and conveyed transparently through the network. At the entrance of the ATM Pipe Switch, the VPI in an incoming cell is translated to an identifier. On the outgoing link, the cell is given a new VPI.

In the case of VC switching, the VPI and VCI of incoming cells are together translated into an internal label. On the outging link, this internal label is re-translated into a VPI and a VCI.

The VPI and/or VCI translation can be regarded as a kind of switching. This "identifier switching" has some similarities with the switching of time-slots performed in ordinary synchronous switches.

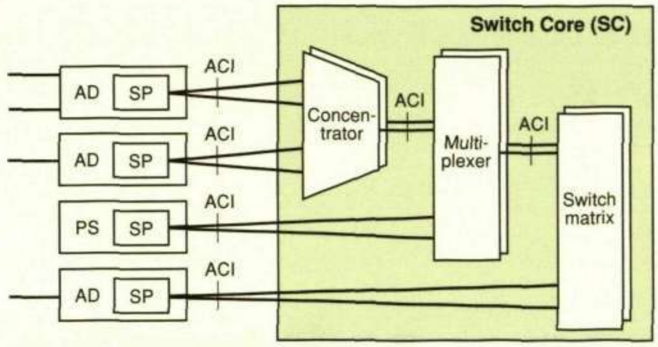


Fig. 5
The Switch Core structure includes concentrators, multiplexers and switch matrixes

ACI	ATM Core Interface
SP	Switch Port
AD	Acess Device
PS	Processor Station

Core. The bit rate and the format of the ATM cells are adapted by the Switch Port to fit the Switch Core. In addition, almost all functions that handle the ATM cell label reside in the Switch Port: VPI/VCI translation, the adding of routing information to the cells, and the discarding of cells with an invalid VPI/VCI. Switch Ports have been designed for various bit rates. Currently, 155 Mbit/s, 622 Mbit/s and 2,4 Gbit/s Switch Ports are developed. The terms VCI and VPI are explained in Box 3.

Virtual Path (VP) switching and Virtual Circuit (VC) switching are simultaneously supported by the Switch Ports, which results in valuable flexibility.

Switch Core

The Switch Core is a space switch which supports both point-to-point and point-to-multipoint connections. Space switching of asynchronous connections implies a demand for large buffer capacity. The Switch Core is equipped with large buffers at the outlets, to cope with variations in the asynchronous cell flow. By using cell-delay and cell-loss priorities in these buffers, the ATM switch can provide for various QOS (Quality Of Service) classes, see Box 1. Connections can be given one of two cell-delay priorities and several levels of cell-loss priorities. Cells belonging to the same connection can be given different cell-loss priorities.

The Switch Core plane can be duplicated to increase the availability of the switch.

The Switch Core is built up of three different types of functional units – concentrators, multiplexers and switching matrices – which can be combined to meet the various capacity requirements for a switching node in a cost-efficient way. Fig. 5 shows a Switch Core containing all three types of units.

Concentrators and Multiplexers

Devices having a lower bit rate than that offered at the Switch Port are connected via concentrators to give better utilisation of the incoming link to the switch matrix.

Multiplexers are used when the switching matrix works at a higher link speed than the Switch Ports; they multiplex the cell flow from a number of ports into a cell stream with the same cell rate as that of the switching matrix.

The Switch Core always contains a switching matrix. Concentrators and multiplexers are added to the Switch Core when concentration or multiplexing of certain – or all – incoming cell streams is a better solution.

The Switching Matrix

The incoming cell streams are passed through to the switching matrix. No switching takes place in the concentrators and multiplexers. The switching matrix is built up as a strictly non-blocking, quadratic, space-switching matrix.

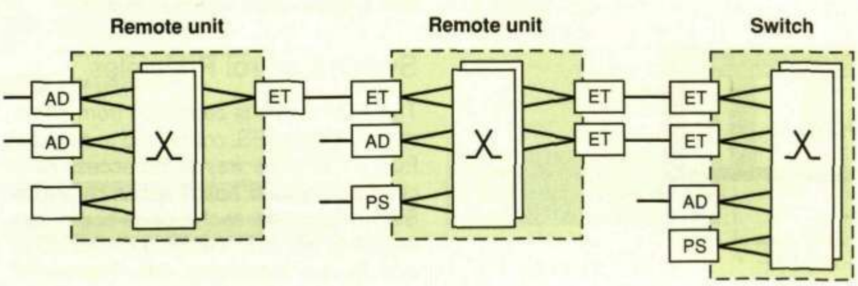
Remote Units

The ATM Switch concept permits cascade connection of remote units. This is accomplished by employing a transmission system with the associated Exchange Terminals, ET, between the remote units and the main network node. The remote units may be small access switches located at a customer's premises. A remote unit may include a Processor Station but can also rely on a Processor Station located in the main network node.

Fig. 6 illustrates remote units connected to a large network node. Depending on the availability requirement placed on the traffic through the remote unit, this latter will be equipped with a single or duplicated Switch Core plane and a single or dupli-

Fig. 6
Remote units connected in cascade to a network node. Exchange Terminals connect the switch and the units to the transmission system. The remote unit may include a Processor Station but can also rely on a processor located in the main network node

AD	Access Device
ET	Exchange Terminal
PS	Processor Station



cated transmission link to the main network node.

Software Architecture

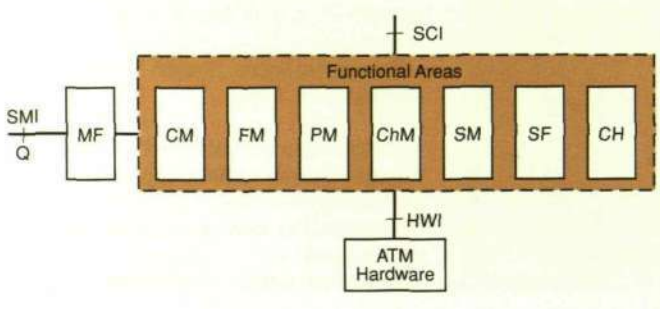
The ATM switch is controlled and supervised by software configured into three categories: one handling traffic management, the second handling operation and maintenance, and the third handling system functions. These three categories are, in turn, divided into seven functional areas, viz:
- Connection Handling (CH)
- Configuration Management (CM)
- Fault Management (FM)
- Performance Management (PM)
- Charging Management (ChM)
- Security Management (SM)
- System Functions (SF)

An essential requirement when designing software for the ATM switch has been the partitioning of it into hardware-dependent and hardware-independent software, and to structure it in an efficient way. Software blocks with well-defined interfaces to other software blocks, and with as little interdependency as possible, are a prerequisite for simplified, systematic upgrading of the switch and for the introduction of new functionality.

Fig. 7 illustrates the functional areas within the ATM switch and the interfaces which encapsulate the software. The Switch Control Interface is an interface towards the application- and service-dependent software. Managed Objects (lines, routes, subscribers, etc) for the ATM switch have been defined and implemented. A Mediation Function (MF) is also under way; designed so as to harmonise with the framework that is currently being created by the CCITT and ETSI.

Fig. 7
Interfaces and functional areas in the ATM switch software

SMI	Switch Management Interface
SCI	Switch Control Interface
HWI	HardWare Interface
MF	Mediation Function
CM	Configuration Management
FM	Fault Management
PM	Performance Management
ChM	Charging Management
SM	Security Management
SF	System Functions
CH	Connection Handling



Redundancy Principles

The ATM Switch concept will be used for different applications – from small business hubs to large public network nodes. Availability requirements will vary with the applications, and the switch must be capable of meeting these requirements in a modular and flexible way.

Network nodes with higher availability are achieved by adding a parallel Switch Core plane.

The redundancy in the Switch Core terminates in the Switch Ports. From each Switch Core plane, the Switch Ports – under faultless conditions – receive a stream of identical ATM cells. Each Switch Port decides which Switch Core plane provides the best quality for the connection in question. It follows that both Switch Core planes are active simultaneously.

This means that the Switch Core has not only redundant planes but also redundant paths. A single fault in the Switch Core will only affect one path in the Switch Core plane and not the whole plane. The availability of the ATM switch will thus be considerably increased.

The redundant Switch Core plane is used when availability requirements are stringent. In other cases the reliability of a single plane is sufficient.

The design of the ATM switch does not apply the principle of duplicating individual access devices and Switch Ports. Instead a robust preselector can be added, allowing a number (n) of physical links to be connected through a larger group (m+n) of access devices, and thus offering n+m' redundancy. Fig. 8 shows an ATM switch with a preselector attached to it.

Switch Control Principles

The ATM switch is controlled from a Processor Station, PS, connected to a Switch Port in the same way as an access device. Each printed board assembly in the Switch Core and each device board connected to the ATM switch is controlled by one Device Processor, DP. The switch

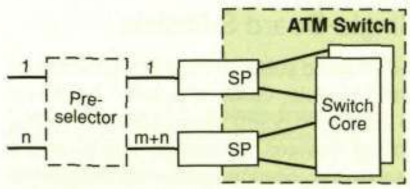


Fig. 8
A preselector connected to the ATM switch

control is hierarchically structured: the PS is superior to the DPs.

Signalling between the PS and the DP uses ATM cells which are transmitted via semi-permanent virtual connections, established over the ATM links. The PS, therefore, need not be physically connected to the switch it controls. When start or restart is initiated, the switch is not configured and the ordinary signalling connections are not established. To enable the PS to reach the DPs during the start/restart procedure, the switch supports a specific signalling connection.

This specific connection is carried by ATM cells, with a pre-determined VPI/VCI and a destination address located in the ATM payload. This signalling connection need not be established; the switch ensures that the cells are routed to their intended destination. The tranmission capacity of the connection is lower than that of ordinary signalling connections.

Clock Principle

Depending on availability requirements, the network node is allocated one, two or three clocks – each on a separate printed board assembly – and an external clock reference.

The clocks are placed in device board positions in the ATM switch. Each clock board distributes one clock signal to each board in the Switch Core, where one of the clocks is selected as active. The Switch Ports receive their clock signals from the Switch Core. In very small ATM switches – those contained in a single subrack, for example – the clocks may be placed on the same printed board assembly as the Switch Core.

External clock references, when needed, can be connected to the system. The frequency of the internal clocks is adjusted to the reference source.

The ATM Pipe Switch

All the components within the Switch Core, i.e. concentrators, multiplexers and switch matrices, are designed according to the Pipe Switch concept. This includes two types of ASIC which in different configurations are used to build Switch Core components of various capacities. The ASICs are named ATM Input Circuit, AIC, and ATM Output Circuit, AOC, respectively.

The AIC terminates the incoming links that carry the ATM cells to the Pipe module, aligns the cell flows from the incoming links, timewise; merges them into a common flow and converts the data format to the Pipe interface format. This format is chosen so that the cell flow can be read in the AOC with a minimum of circuit complexity. The AOC performs switching and buffering and composes the cell flow to be sent on the outgoing links from the Pipe module. A memory, embedded in AOC to cater for the buffering towards the outgoing links, enables each AOC to handle four outgoing links. Several AOCs can be connected in series to multiply the number of outputs from a Pipe module.

The Pipe module always contains one AIC, whereas the number of AOCs depends on the number of Pipe module outputs required – up to the limit set by the size of the printed board assembly.

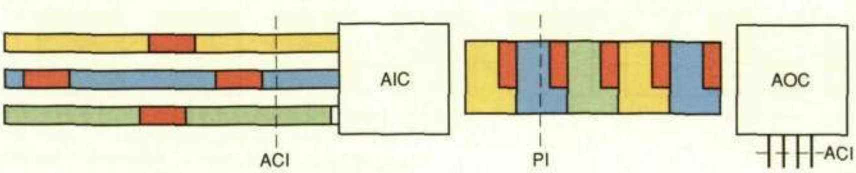
The AIC terminates 32 incoming links, each carrying a 155 Mbit/s serial stream of ATM cells to the ACI interface. These 32 links can be arranged to handle data flows corresponding to 32 links at 155 Mbit/s, 8 links at 622 Mbit/s or 2 links at 2,5 Gbit/s. The AIC converts the incoming cell flows to a common flow on the Pipe interface to an AOC circuit, Fig. 9. The cell stream on the Pipe interface is transparent through the AOC, to make it possible to series-connect an unlimited number of AOCs. The AOC can copy cells from the Pipe to any of the outgoing links on the device.

Space switching is performed by comparing the routing information inserted in the cells by the Switch Ports with entries in an output-oriented routing table in the AOC. When the routing information and an entry in the table match, the cell is copied to the

Fig. 9
AIC performs two basic functions: transformation of the incoming cell streams on the ACIs to the Pipe interface, and cell alignment

AIC ATM Input Circuit
AOC ATM Output Circuit
ACI ATM Core Interface
PI Internal Pipe-Pipe Interface

ATM cell header



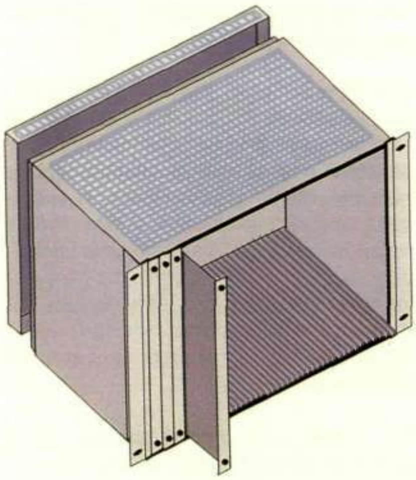


Fig. 10
Mechanical design of the GBM
Widht 450 mm
Height 450 mm
Depth 200 mm

corresponding output buffer. For point-to-multipoint connections, the routing table has an entry for each of the outputs from which the cell is to be sent.

The cell buffering function uses an integrated static-memory pool common to all four outputs. Cell-loss and cell-delay priorities are handled when the cell is stored in the buffer. There is also functions to notify the user about high buffer loads by means of the standardised Explicit Forward Congestion Indication (EFCI) mechanism, Box 1. From the buffers, the cells are transferred to the output links. PROMs are used to set the cell rate to be used at the outputs so that it corresponds to bit rates 4x155 Mbit/s, 4x622 Mbit/s or 1x2,5 Gbit/s.

The outputs from the AOC use the same ACI format as the inputs to the AIC. In this way, several Pipe components can be connected to form different Switch Core configurations.

Single-Board Switches

A dedicated subrack called Generic Broad band Module, GBM, is used to house up to 20 printed board assemblies containing access devices, as well as clocks and a Processor Station with its integrated Switch Ports, Fig. 10. In the GBM there are also two dedicated positions for single-board switch matrixes, one position for each redundant plane. A small single-board switch designed to fit into the GBM subrack can handle 20 links at 155 Mbit/s – one for each access board in the GBM, Fig. 11. Interconnected GBMs in the Switch Core are controlled by the same clocks.

For applications where the need for total bandwidth capacity is limited, a switch matrix with a concentrating Pipe, Fig. 12, is used. In this Pipe, one AOC output is branched to several Switch Ports, which increases the number of outputs. All these Switch Ports receive the same cell flow, and each of them discards all cells with a VPI/VCI that is invalid for the Port.

Single-Board Concentrator

For accesses which, together, only demand a fraction of the total link capacity of 155 Mbit/s, a concentrator is employed to utilise the switch matrix capacity more efficiently. The concentrator transforms 20 incoming links, each at 155 Mbit/s, to one link at 155 Mbit/s towards the switch matrix. The concentrator is built on a single board that can replace a single-board switch matrix in a GBM.

Single-Board Multiplexer

A multiplexer is used to enable several links at 155 Mbit/s to access a switch ma-

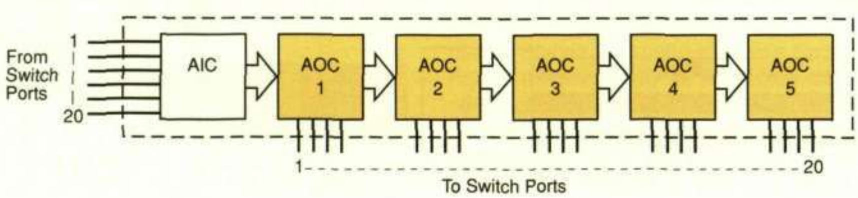


Fig. 11
A Pipe configuration for a single-board switch for 20 links at 155 Mbit/s. Note that only 20 of the 32 inputs are used

AIC ATM Input Circuit
AOC ATM Output Circuit

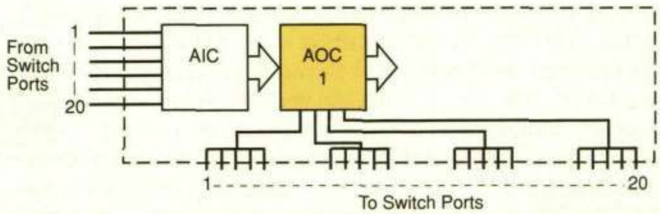


Fig. 12
A Pipe configuration for a single-board switch for 20 links at 155 Mbit/s with reduced total throughput. Each output from the AOC is connected to several Switch Ports through a passive branching function

trix operating at 2,5 Gbit/s. The multiplexer – built on a single board that can replace a single-board switch matrix in a GBM – is designed to handle twenty 155 Mbit/s links, which are multiplexed to one 2,5 Gbit/s link. Two Pipes are used, one for multiplexing and one for demultiplexing, Fig. 13. Note that although the multiplexer can connect 20 links at 155 Mbit/s the total bandwidth utilisation should not exceed 2.5 Gbit/s, (16x155 Mbit/s) to avoid blocking on the 2.5 Gbit/s link.

ATM Switch Module 20G

The ASM 20G contains a switch matrix for eight 2.5 Gbit/s links, which gives a total throughput of 20 Gbit/s. ASM 20G is built in one subrack and uses two types of printed board assemblies with pipes. The actual switching is performed on four switch boards, each with a four-link switch matrix for 2.5 Gbit/s, Fig. 14. The other board type, split-and-merge board, is used to obtain a quadratic expansion of the switch boards by splitting the incoming links to a row of two switch boards and merging the outgoing links from a column of two switch boards, Fig 15.

ATM Switch Modules ASM 40G and AMS 80G

ATM switches with capacities exceeding 20 Gbit/s are built up of interconnected ATM 20G units. The ATM Switch Module 40G uses four ASM 20G in a quadratic expansion. This expansion is effected through split-and-merge boards in the same way as the switch boards are expanded internally in the ASM 20G. An ASM 40G can handle 16 links at 2.5 Gbit/s or, with multiplexers, 256 links at 155 Mbit/s.

In a recursive manner, a quadratic expansion can be effected by using the ASM 40G to form an ASM 80G with a total throughput of 80 Gbit/s, handling 512 fully utilised 155 Mbit/s links.

Conclusions

The ATM switch offers an asynchronous, connection-oriented, cell bearer service. This basic function can be enhanced by additional software and hardware, to implement a number of high-level bearer services. A valuable characteristic of the ATM switch is that it allows different bearer ser-

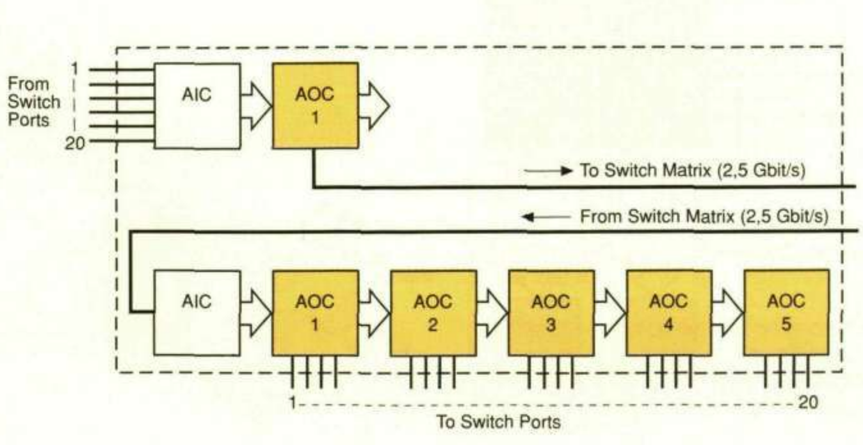


Fig. 13
A two-pipe configuration for a multiplexer. 20 links at 155 Mbit/s are multiplexed to one link at 2.5 Gbit/s. The top pipe is used for multiplexing and the bottom pipe for demultiplexing

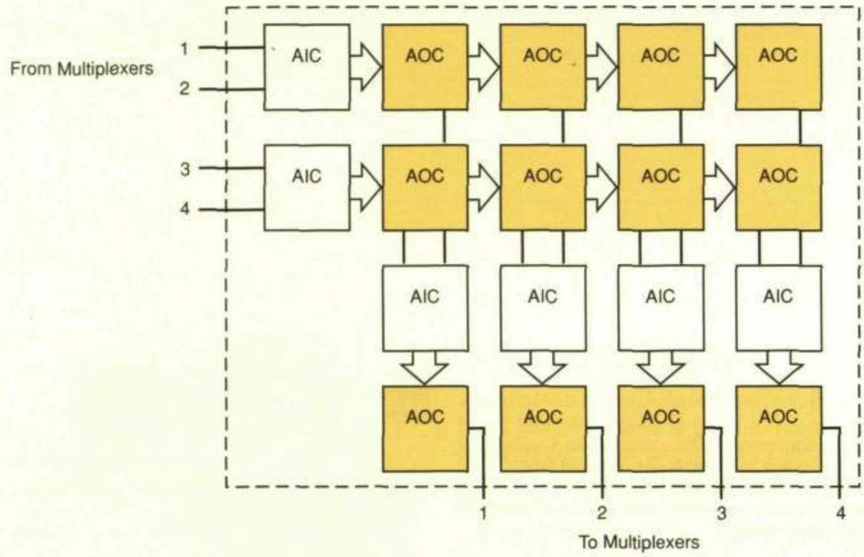


Fig. 14
Pipe configuration for a Switch Board with a 4 x 4 switch matrix operating at 2.5 Gbit/s

vices to share its switching resources simultaneously.

The strategy for the development of the ATM switch has been that of strictly separating the switch concept from a specific hardware implementation. The ATM Switch concept is characterised by

- a small number of stable interfaces for data transport, control and operation
- strict separation of ATM switch functions and broadband application functions
- software operating on an abstract model of the hardware.

The ATM Switch concept can be used for private and public applications, for centralised and remote nodes and nodes connected in cascade.

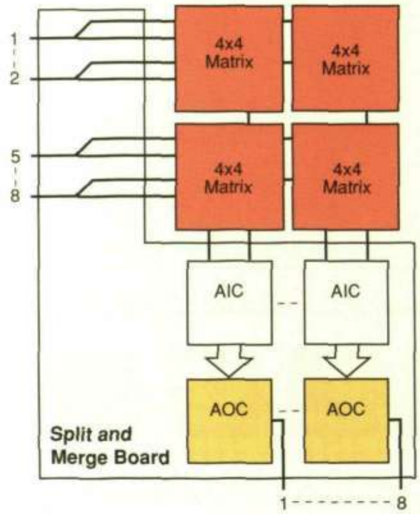
A prototype version of the ATM Pipe Switch and a number of broadband applications for the switch are now under test.

The Pipe Switch is a non-blocking switch with up to 80 Gbit/s throughput, characterised by a modular structure based on two full custom integrated circuits. The minimum configuration, consisting of only two components, will be a useful tool for cost-effective solutions in small-scale applications, e.g. private network hubs.

Access lines for 155 Mbit/s, 622 Mbit/s and 2.5 Gbit/s can be connected to the ATM Pipe Switch and provide point-to-point and point-to-multipoint connections with support of cell-delay and cell-loss priorities.

The switch has a triplicated clock system and can be equipped with a single or duplicated Switch Core plane.

Fig. 15
An 8 x 8 switch matrix with a throughput of 20 Gbit/s is built up by four 4 x 4 matrix boards and a split-and-merge board



References

1 CCITT Recommendation: I.150 B-ISDN Asynchronous Transfer Mode Functional Characteristics.

RMS – an AXE 10 System for Measurement of Transmission Quality on Telephone Circuits

Francisco J. Golderos

Among telecommunications network operators there is a great demand for automated operation and maintenance systems. The use of relevant and accepted parameters for verifying the quality of circuits is also becoming more and more important. To meet these requirements, Ericsson has developed an AXE 10 subsystem for measuring transmission quality on digital and mixed analog/digital telephone circuits.
The author describes the functions, structure and capacity of the system and outlines some of its configurations.

FRANCISCO J. GOLDEROS
Ericsson, S.A.
Spain

Monitoring the transmission quality on trunk circuits constitutes a fundamental part of the operation and maintenance functions in a telecommunications network. The final network user requests a service level that satisfies his needs. The network operator, on his part, continuously needs information about the condition of the network to be able to maintain the requested quality and remedy any deficiencies. Since data communication represents an ever-growing portion of the traffic in telecommunications networks, the demand for high availability of the services provided to the customer is on the increase.

In the future, the demand for high-quality circuits will be even greater than today, and the use of relevant and accepted parameters will therefore be important in measurements. The quality parameter currently used for digital circuits is Bit Error Ratio (BER), with alarm thresholds at (normally) 10^{-3} and 10^{-6}, which is not good enough for data traffic. The measurement of transmission quality, both on analog and digital lines circuits, is therefore of great importance for the maintenance and supervision of the network.

General Description of RMS

Ericsson has developed a real-time system – Remote Measurements Subsystem (RMS) – for digital and mixed analog/digital trunk circuits. The RMS, which is typically used to perform measurements and routine tests and to localise faults, can test analog trunks and 24- and 32-channel PCM systems.

Operator commands are used to initiate tests, or for the preparation of tests to be initiated later on. The test position is equipped with a terminal – a PC or a workstation – and one or more telephone sets. The RMS can be operated from a remote location.

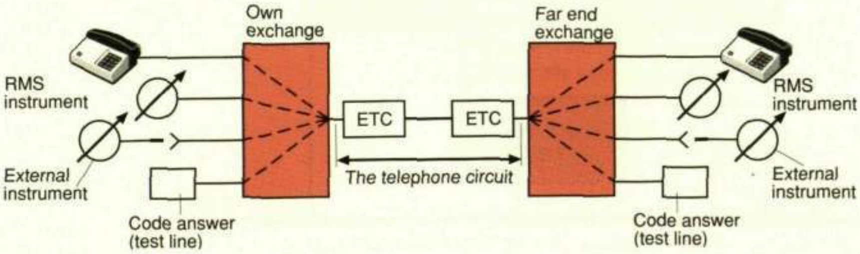
From its first introduction, the system has been improved to facilitate handling and – by providing new functions – to meet new demands from CCITT and different markets.

The following test functions are supported by the RMS:
- Level measurement
- Noise measurement
- Gain Slope Level measurement
- Echo Return Loss measurement
- Singing Return Loss measurement
- Transmission of continuous tones
- Transmission of bit patterns
- Bit Error Ratio measurement.

All these functions are described in detail in the following.

The circuits to be tested can be accessed with or without signalling. Signalling here refers to the signals that set up connections. If signalling is used, the termination

Fig. 1
Measurements with RMS
Code answers in the remote exchange facilitate automatic measurements. In the other two alternatives, test calls require assistance by an operator in the far-end exchange. The RMS at the near end sets up speech connections through the circuits to be tested. The RM instrument at the far end is used for terminating test calls. When an RMS is available at the far end, it can connect the RM instrument to the test call, by command from the operator. When no RMS is available at the remote exchange, the operator must connect the instrument manually



in the remote exchange associated with the B-number is accessed through the circuit under test. (The B-number can indicate an access point in the Subscriber Stage or the Group Switch). When no signalling is used, the RMS blocks the circuit to be tested and, consequently, knows its identity.

The RMS can interwork with remote exchanges using different equipment at the far end:
- Test lines (Code answers)
- RM instruments
- Manually controlled instruments.

Fig. 1 gives an overview of the different possibilities. The RMS is equipped with instruments for measurement of transmission quality, and transmitters which can send tones and different bit patterns. The RMS is connected to the trunks to be tested through the AXE Group Switch.

Different equipments can be connected at the remote exchange. When Code answers are used in the remote exchange, fully automatic measurements are possible. The only intervention needed from the operator at the near end is to define and initiate measurements. The Code answer provides a fixed tone-pause message, a digital loop-around, a quiet termination, or a combination thereof. No technician is needed at the remote end, since the RMS knows what response the Code answer gives. If the RMS sends signals, they are looped around by the Code answer and thereby enabling the RMS to measure. Alternatively, the Code answer sends the signals backwards from the far end. The 'quiet termination' alternative is used for noise or echo measurements, for example.

In the two other alternatives in Fig. 1 calls are made to an operator-attended test position in the far-end exchange. The RMS sets up speech connections between operators located at different exchanges through the circuits to be tested. In cases where an RMS is available in the remote exchange, the operator only enters a command that switches the connection to the relevant RM instrument. When no RMS is available, the operator must connect the instrument manually.

System Functions

RMS contains fifteen different functions which fall into three main functional areas, fig. 2:
- User Functions
- Administration Functions
- Measurement Functions.

User Functions
User functions
- act as interfaces between the operator and the system
- control the measurements
- have access to all the Administration and Measurement functions in the system and invoke those that are specified for a particular measurement.

The types of measurement to be made and/or signals to be transmitted, as well as the test objects in a particular test, are defined in a User function. This is done with commands by the operator.

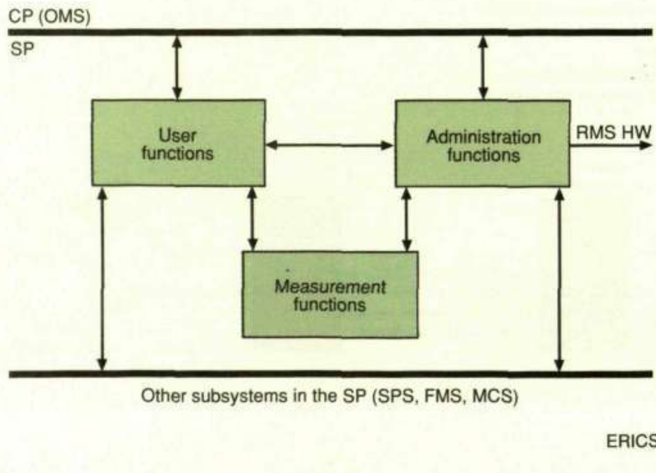


Fig. 2
RMS functions fall into three main function areas:
- **User Functions control the measurement performance of the RMS**
- **Administration Functions handle the RM instruments and provide the User Functions with special facilities**
- **Measurements Functions support the different types of measurement (level, noise, bit error ratio, etc) and tone sending**

There are four User functions:
- The Remote Measurements on Telephone Circuits function
- The Delayed Measurements function
- The Timetable-Controlled Measurements function
- The CP Located Users function.

The first three of these functions have a great deal in common.
- They offer the operator a set of commands for the handling of measurements
- They accept all data related to a particular measurement to be performed, e.g. the lines to be tested and the types of measurement or signals to be transmitted. All measurement functions that the RMS supports, as well as preprogrammed sequences of measurements (interwork with test lines), can be activated. The test objects can be both digital and mixed digital/analog circuits
- They make it possible to establish a path through the AXE Group Switch, between the seized RM instrument and the circuit to be tested. This also includes the possibility of setting up connections – for listening and speech, through the circuits to be tested – between operators located at different exchanges.

The Remote Measurements on Telephone Circuits function
The function 'Remote Measurements on Telephone Circuits' is used when the test is to start immediately.

This function is based on the 'Immediate Measurement Session' concept, which is created with an initiating command each time an immediate measurement/transmission is started. If several measurements are required in parallel, an Immediate Measurement Session is initiated for each of them.

A single command is sufficient to start automatic measurement of all circuits belonging to a route or up to 32 individual circuits. A defined measurement can also be automatically repeated on a circuit (continuous measurement) until the operator interrupts the process.

The Delayed Measurements function
The function `Delayed Measurements' is used when the operator wants measurements to be performed at a later time.

This function is based on the 'Delayed Measurement Session' concept. A Delayed Measurement Session is created with commands by the operator. In addition to the definition of types of measurement and test object, the date and time of starting the test are also given. The session will remain active until the measurements have been performed. If several measurements are required in parallel, a Delayed Measurement Session is initiated for each of them.

Up to 128 Delayed Sessions can be defined at the same time. This maximum value has been considered as reasonable. It can be increased in the future at operators' request.

The Timetable-Controlled Measurements function
The function 'Timetable-Controlled Measurements' is used when the operator wants to schedule routine measurements.

A set of measurements, which is defined in the RMS Timetable, is cyclically repeated. The RMS Timetable covers a period of up to 84 days and consists of a number of cells, one for each hour during each day. Several types of measurement can be scheduled for different test objects (routes) within the same cell.

Tests are started automatically at the time they are scheduled (in one of the timetable cells), and continue until all available circuits have been tested or the time for the cell elapses. In routes with a large number of circuits, measurement data need to be repeated in adjoining cells so that all circuits can be tested.

To meet the needs of different network operators, various lengths of the timetable can be chosen. When a cycle is completed it will start again from the beginning after a period of time. This variable-length intercycle period allows the operator to synchronise the dates when tests are regularly started, Fig. 3.

The timetable based on CCITT Recommendation M.605 is of special interest. This is a fixed-length timetable covering 56 days and starting on the first monday of all odd months.

A set of commands is provided for the handling of data in the timetable and to activate/deactivate the RMS Timetable on a specific date and at a specific hour.

The Central Processor Located Users function
The RMS uses part of the Terminating Test Call function that belongs to the Operation and Maintenance Subsystem (OMS) in AXE 10. In the RMS, this part is referred to as the CP Located Users function. It accepts orders from the CP to switch between operator and RM instruments. The connection path between the test objects and the instruments is provided by the Terminating Test Call function in OMS.

Administration functions
Administration functions control the RM instruments and provide the User functions with facilities to be used during the measurement. There are four Administration functions:
- Instrument Administration
- Standard Adjustments of Instrument Data
- Programmed Interwork with Test Lines
- B-Number Translation.

The Instrument Administration function
The Instrument Administration function allows the operator to connect and disconnect, block and deblock, and test the RM instruments by means of commands. The function selects the relevant instrument and controls it during the measurement. It also handles alarms that indicate any fault in the RMS hardware and files.

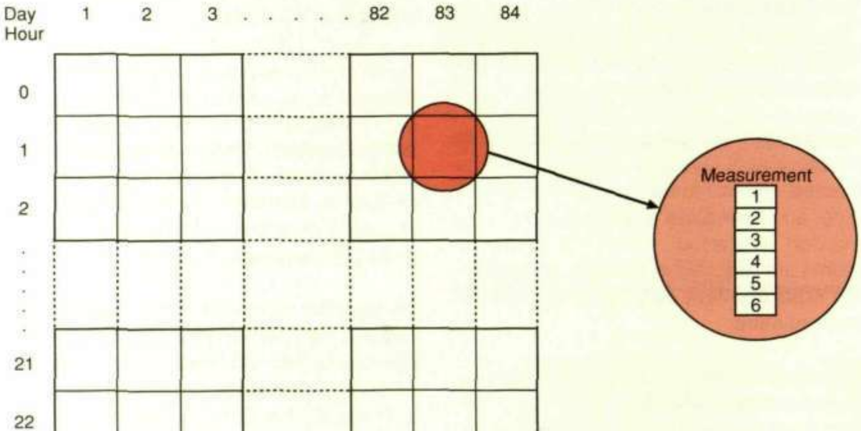
The Instrument Administration function, interworking with the OMS in AXE 10, controls the Switching Network Terminals (SNT), which administer the connection of RM instruments to the Group Switch. OMS acts as SNT-owner of the PCD-D (Pulse Code Modulation Device-Digital) multiplexer for the RM instruments (see below).

The Standard Adjustments of Instrument Data function
The function Standard Adjustments of Instrument Data enables the operator to modify the standard adjustment normally used during a measuring process. The standard adjustments are system default values that the User functions normally use during the RM measurements.

The Programmed Interwork with Test Lines function
The function 'Programmed Interwork with Test Lines' makes it possible for the operator to set up sequences of measurements in which the RMS interworks with test lines (Code answer). A sequence of measurements is a set of single measurements of level, noise and echo return loss, for example. The Code answer equipments have one or more of the following capabilities: tone sending, quiet termination and loop-around. An example of a measurement sequence is:
- Start
- Measure level
- Wait 3 seconds
- Measure noise
- Wait for Clear-Back indication
- Measure echo return loss
- Stop.

The following types of test line can for example be used:
- type O.11 according to CCITT Rec. O.11
- CANSC Code Answer, code 100
- CANS Code Answer, code 102
- TONE Test Tone
- TONED Test Tone Digital

Fig. 3
The RMS timetable covers up to 84 days. Each hour corresponds to one cell. One cell can support up to six preprogrammed measurements. Thus, more than 12,000 tests can be performed during one cycle when the timetable is active. The inter-cycle time can be set by the network operator to fit the routine test schedule



Box 1

Error measurements and bit pattern characteristics

Bit patterns according to CCITT Rec. O.152:
- a Pseudo-random pattern with a length of 2047 bit
- b Pseudo-random pattern with a maximum of 7 consecutive '0'-bit
- c Fixed pattern of ones (...1111...)
- d Fixed pattern of alternating ones and zeroes (...0101...)

Bit patterns with errors:
Bit patterns *a* to *d* in which the operator can introduce a number of bit errors. The time with bit errors and the error- free time are also specified by the operator

Fixed test patterns:
A well-known result is produced in the following situations:
- when three fixed test patterns are compared with pattern *a*
- when another fixed test pattern is compared with pattern *b*

User-defined test pattern:
The operator can define his own test patterns by means of commands. The test patterns can have a length of one to four bytes

Error measurements:
- Bit Error Ratio (BER) for the available time
- Block Error Ratio (BLER) for the available time
- Errored Seconds (ES %)
- Severely Errored Seconds (SES %)
- Degraded Minutes (DM %)
- Unavailable Seconds (UAS %)
- Number of resynchronisations
- Number of detected SLIPs

Up to 30 different sequences of up to eight single measurements can be defined by means of commands.

The B-Number Translation function
The B-Number Translation function allows the operator to translate B-numbers to mnemonic names (B-names) to be used by the preprogrammed measurements (delayed or timetable-controlled). This makes it easy for the operator to specify the access number to terminals in the cooperating exchanges. He needs not remember the digit sequence of the B-number. Furthermore, if a B-party access number is removed or changed, the preprogrammed measurements that use this number need not be modified; the function allows the B-number associated with the B-name to be changed.

Measurement functions
A Measurement function can perform a measurement or send tones or bit patterns as requested by the User functions.

The Transmission of Tones function
The Transmission of Tones function makes it possible to send tones of a specified frequency and level on mixed analog/digital circuits in compliance with CCITT's Recommendation O.11.

A-law and μ-law can be used. The frequency of the tones ranges from 50 to 3950 Hz in steps of 1 Hz. To avoid harmonic distortion, 0.31873 Hz is added to the transmitted frequency.

The level of tones to be sent ranges from –60 dBm0 (A-law) and –70 dBm0 (μ-law) to +3 dBm0, in steps of 0.1 dB.

The Level Measurement function
The Level Measurement function measures the level and the frequency of the strongest tone received over mixed analog/digital circuits in compliance with CCITT's Recommendation O.22. The tone to be measured can be generated in the interworking exchange or by the same instrument as the one used for the measurement.

The level of the signal can be calculated as a root mean square or as an average value. The frequency of the measured signal is detected in a range from 0 to 4000 Hz, with a resolution of 1 Hz. The level ranges from –100 to +6.18 dBm0, with a resolution of 0.01 dB.

The Noise Measurement function
The Noise Measurement function allows measurement of idle channel noise and noise with a holding tone on mixed analog/digital circuits in compliance with CCITT's Recommendation O.22. When the noise is measured, the holding tone is rejected with a band stop filter (notch filter). The holding tone can be generated in the interworking exchange or by the instrument used for measuring.

A weighting filter evaluates the effects of noise corresponding to the annoyance it causes a "typical" subscriber to standard telephone circuits. The noise can be weighted by means of a C-message weighting filter (U.S.A) or a psophometric filter (CCITT Rec. O.41).

The noise is calculated as a root mean square value. Noise values range from –100 to +6.18 dBm0, with a resolution of 0.01 dB.

The Gain Slope Level Measurement function
The Gain Slope Level Measurement function measures the attenuation distortion on mixed analog/digital circuits in compliance with Bell Technical Reference PUB 41,009. A sequence of three tones (1004, 404 and 2804 Hz at –16 dBm0) is sent automatically and looped back at the remote exchange.

The Echo Return Loss Measurement function
The Echo Return Loss (ERL) Measurement function measures the attenuation of a band-limited noise signal (560–1965 Hz at –10 dBm0) on mixed analog/digital circuits, in compliance with Bell Technical Reference PUB 41,009. The function presents the difference in amplitude between the injected test signal and that portion of the initial signal which the hybrid circuit fails to suppress. The attenuation range is from –16.18 to +90 dB with a resolution of 0.1 dB.

The Singing Return Loss Measurement function
The Singing Return Loss (SRL) Measurement function measures the attenuation of two band-limited noise signals (260–500 Hz for SRL low and 2200–3400 Hz for SRL

high at –10 dBm0) on mixed analog/digital circuits, in compliance with Bell Technical Reference PUB 41,009. Experience has shown that there are two frequency ranges in which a circuit can oscillate, and for this reason two different measurements must be made in order to verify in which of these ranges the critical frequency lies. The function presents the difference in amplitude between the injected test signals (one at a time) and the respective reflected signals.

The Bit Error Ratio and Transmission of Bit Patterns function
This function makes it possible to perform Bit Error Ratio (BER) measurements and to send bit patterns on digital circuits, in compliance with CCITT Recommendations G.703, G.821 and O.152. With the RMS, the function can be used on 24-channel digital circuits that employ bit stealing, which means that it can be used on PCM systems with 56 or 64 kbit/s. The types of bit pattern that can be generated by the function, as well as the technical characteristics of the function, are described in Box 1.

RMS configuration in AXE

Some important considerations in the design of the RMS were:
- to relieve the central processor (CP) of operation and maintenance functions with only moderate real-time demands
- to automate diagnostic and routine tests as far as possible, in order to relieve the operator from manual actions
- to make it easy to introduce new facilities and functions without affecting traffic in progress
- to make it possible to control the measurement system from a remote maintenance centre
- to increase the reliability and safety of the system by duplicating the measurement system without affecting traffic capacity.

The listed goals were achieved by implementing the RMS in a support processor (SP). The SP is a general-purpose computer system (APN 167); the software consists of a number of application programs. An SP can be considered as a *front-end* computer for the AXE 10 control system.

The SPs are interconnected to form Support Processor Groups (SPG). An SPG can consist of several support processors, but only two SPs are considered in the standard configuration. Each SP is a node within the SPG. In non-duplicated systems, a single SP constitutes an SPG; in duplicated systems, two support processors are interconnected through the Inter-Computer Bus (ICB). One node is then active, whereas the other is on stand-by, ready to take over the control if the active node fails. Each SP node is connected to the CP through the RP bus of AXE.

The I/O-system IOG 11 of AXE 10 is also based on the SP concept. It is always installed in the SPG 0, Fig. 4. The RMS can be installed in its own SPG or sharing an SPG with IOG 11. The former option is

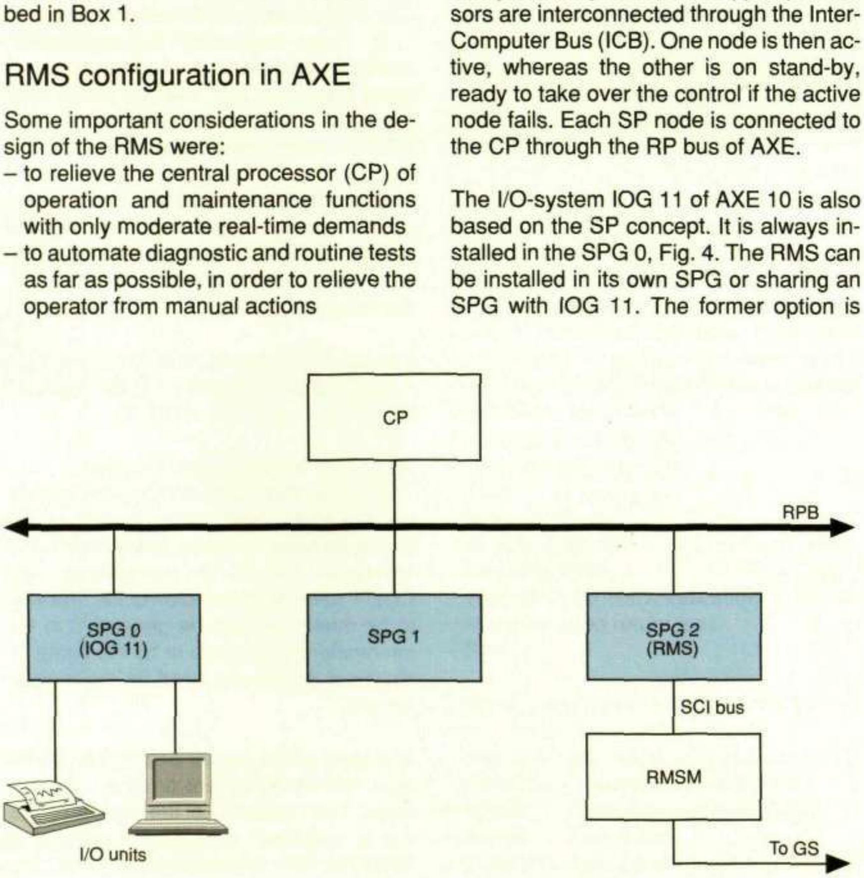


Fig. 4
Typical SPG configuration. The RMS located in the Support Processor Group 2

SPG	Support Processor Group
CP	Central Processor
IOG 11	/O Group 11
RMSM	Remote Measurements Subsystem Magazine
RPB	Regional Processor Bus
SCSI	Small Computer Standard Interface

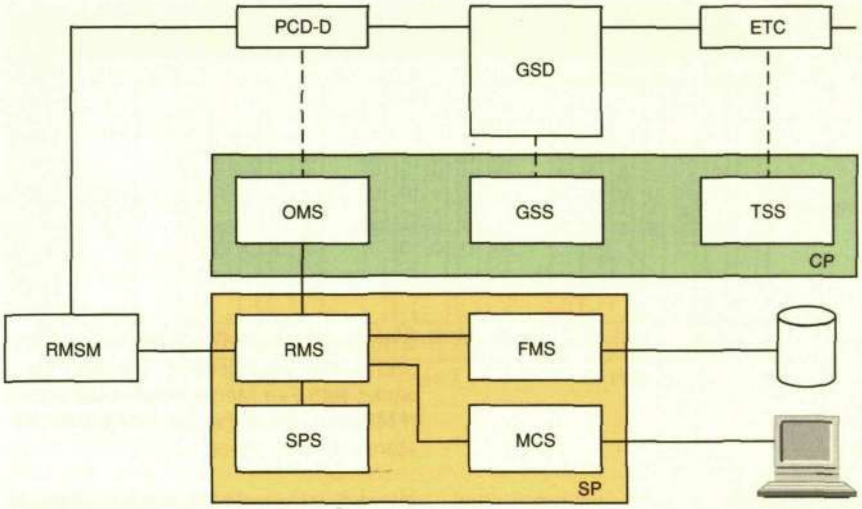


Fig. 5
Configuration of RMS in AXE

ETC	Exchange Terminal Circuit
GSD	Group Switch Device
PCD-D	Pulse Code Modulation Device-Digital
OMS	Operation and Maintenance Subsystem
GSS	Group Switch Subsystem
TSS	Trunk and Signalling Subsystem
SPS	Support Processor Subsystem
FMS	File Management Subsystem
MCS	Man Machine Communication Subsystem
RMS	Remote Measurements Subsystem
RMSM	Remote Measurements Subsystem Magazine

used in large exchanges with a large number of circuits to supervise. Fig. 4 gives an overview of the SPG configuration when the RMS is located in the SPG 2.[1]

The RMS is part of the Switching System APT. It is implemented in both software, executed by the Support Processor, and hardware.

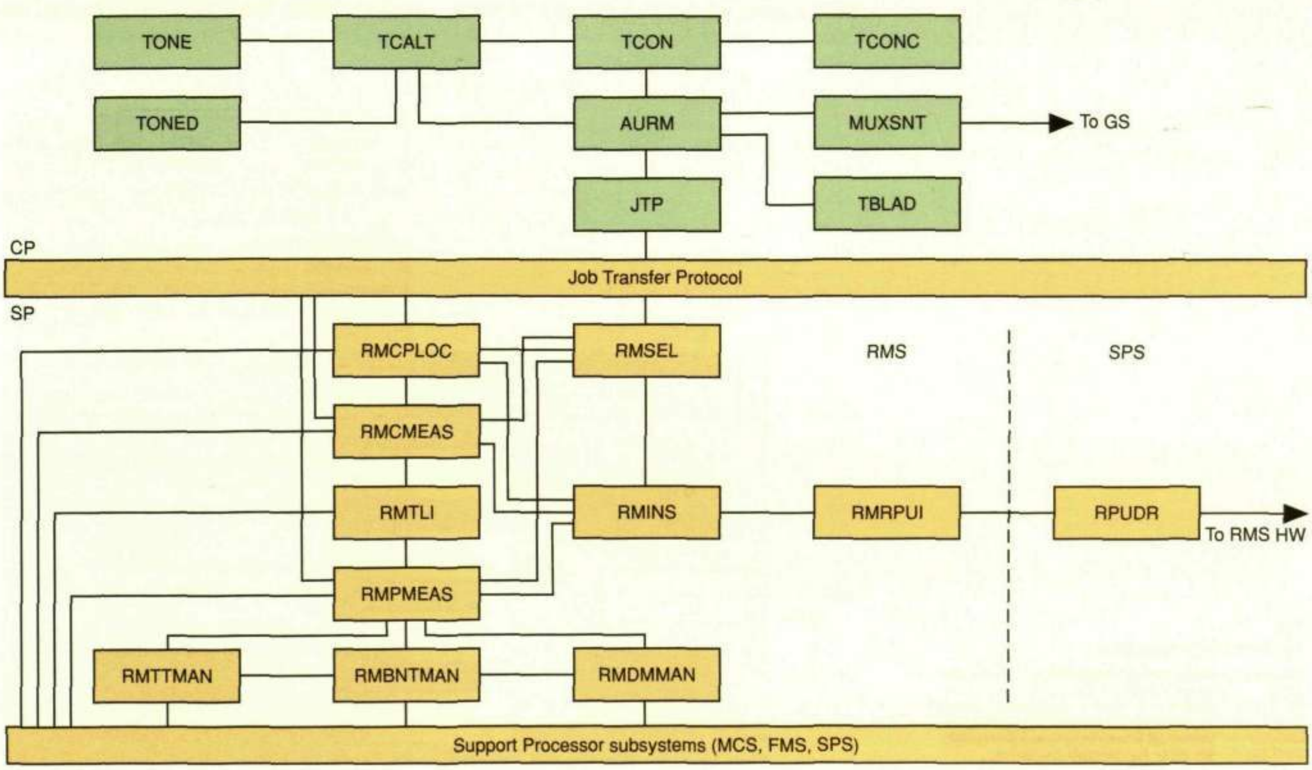
Fig. 5 shows a general configuration of the RMS in AXE. The RMS interworks primarily with the Operation and Maintenance Subsystem (OMS) to set up connections – via the Group Switch (GS) – between RM instruments and test objects in the Trunk and Signalling Subsystem (TSS). Test objects (telephone circuits) that the RMS can measure are Outgoing Trunks (OT), Incoming Trunks (IT) or Bothway Trunks (BT).

The RMS uses most of the basic functions that the Support Processor Subsystem (SPS) provides. For example, the programming language ERIPASCAL based on the real-time operative system ERIOS is used for the RMS. SPS provides the communication function between CP and SP by means of the Job Transfer Protocol (JTP). This protocol is used by the RMS in its interwork with OMS.

The Man-Machine Communication System (MCS) provides the RMS with functions for communication with the operator

Fig. 6
Block structure of the RMS and its interworking with AXE CP function blocks

TONE	Test Tone
TONED	Test Tone Digital
TCALT	Terminating Test Call
TCON	Test Connections
TCONC	Test Position Administration
AURM	Adaption Unit for Remote Measurements
MUXSNT	Switching Network Terminal for Test Equipment
TBLAD	Test-Blocking Administration
JTP	Job Transfer Protocol
RMTTMAN	RM Timetable Manager
RMCPLOC	RM CP Located Users
RMCMEAS	RM Command Controlled Measurements
RMTLI	RM Test Line Interwork
RMPMEAS	RM Preprogrammed Measurements
RMBNTMAN	RM B-Number Translation Manager
RMSEL	RM Instrument Selection
RMINS	RM Instrument Handling
RMDMMAN	RM Delayed Measurements Manager
RMRPUI	RM Regional Processor Unit Interface
RPUDR	Regional Processor Unit Driver



terminals connected to SP or to operation and maintenance centres. The RMS interworks with File Management Subsystem (FMS) to manage the dedicated RMS files stored on hard disks.

Interaction between the function blocks in which the functions are implemented and the interface with other subsystems are illustrated in Fig. 6.

Hardware Structure

The RMS is based on two types of instrument: RMBER (RM Bit Error Ratio) and RMUNI (RM Universal Instrument).

The RMBER instrument is implemented in the AXE standard printed board assembly BER-SC (Bit Error Ratio-Small Computer). It is used for BER measurements and for transmission of Bit Patterns on 64 kbit/s PCM systems without bit stealing.

The RMUNI is implemented in the AXE standard printed board assembly UNI-SC (Universal Instrument-Small Computer). RMUNI is used for transmission of tones and Bit Patterns and to perform the analog-type measurements that RMS supports. Furthermore, RMUNI can perform BER measurements on 56 and 64 kbit/s PCM systems with or without bit stealing. UNI-SC is based on the Fast Fourier Transform (FFT).

The instruments are connected to AXE by the Support Processor Subsystem (SPS) hardware, Fig. 7. A measurement unit consists of four instrument boards and one Regional Processor Unit-Small Computer (RPU-SC) that connects the RMS instruments to the system through the External Bus Adapter-Small Computer (EBA-SC) board.

BER-SC consists of only hardware which is administered by one microprocessor in RPU-SC. An UNI-SC instrument is built-up with two microprocessors, one for transmitting and one for receiving (measuring).

The RMS hardware is housed in the Remote Measurements Subsystem Magazine (RMSM). The RMSM has several possible configurations, depending on the requirements of the exchange. The magazine can be equipped with one or two measurement units, two power supplies (one for the instruments and one for the interface), eight or sixteen Pulse Code Modulation channels and a mixture of instruments within the measurement unit. An

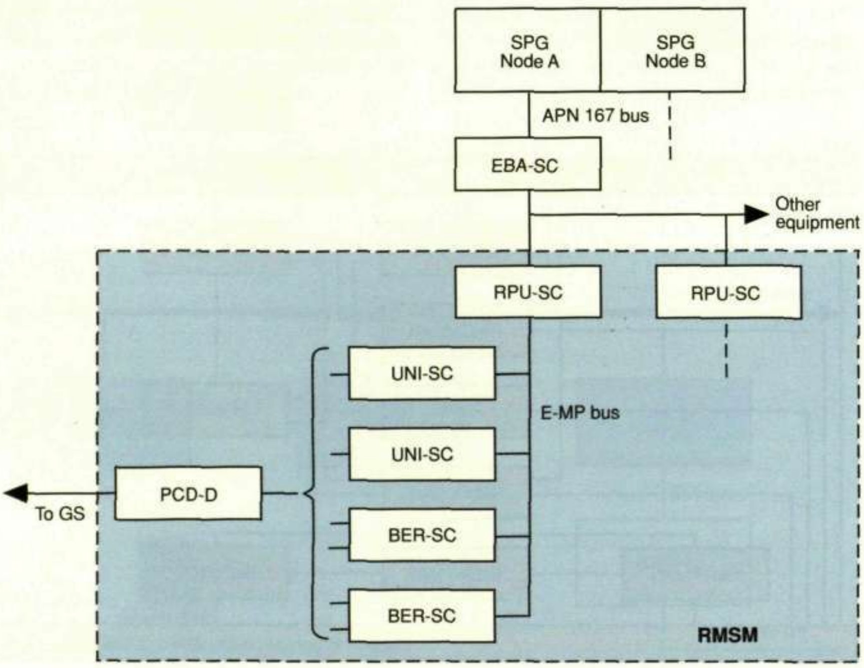
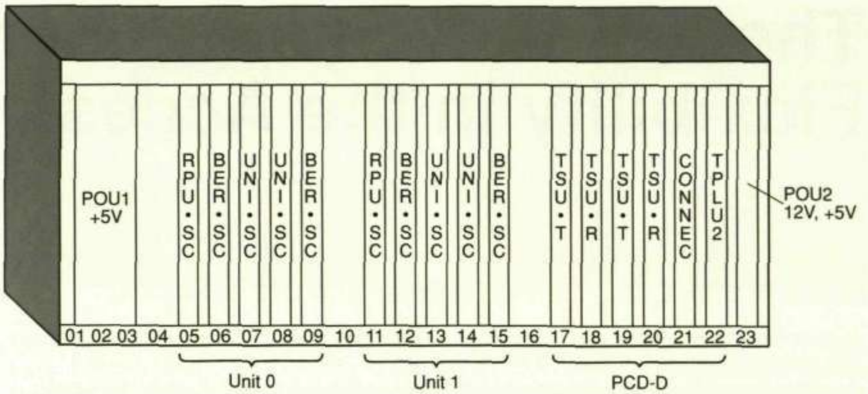


Fig. 7
RMS hardware structure

EBA-SC	External Bus Adapter-Small Computer
SCSI	Small Computer Standard Interface
RPU-SC	Regional Processor Unit-Small Computer
PCD-D	Pulse Code Modulation Device-Digital
UNI-SC	Universal Instrument-Small Computer
BER-SC	Bit Error Ratio-Small Computer

Fig. 8
Example of layout of a Remote Measurements Subsystem Magazine (RMSM). The instrument distribution corresponds to BFD 324,607/14.

POU	DC/DC Converter
RPU-SC	Regional Processor Unit-Small Computer
BER-SC	Bit Error Ratio-Small Computer
UNI-SC	Universal Instrument-Small Computer
TSU-T	Timeslot Unit Transmit
TSU-R	Timeslot Unit Receive
CONNEC	Connection Unit
TPLU2	Timing and Plane Selection Unit



example of the RMSM layout is shown in Fig. 8.

Summary

For network operators, the RMS will ensure improved possibilities of centralisation and automation of maintenance functions. Since the first version of the RMS was released, the RMS user interface has been improved, and new functions have been added to meet the demands of the market. At present, the introduction of new functions is under study. For example, introduction of CCITT Recommendation O.27 (In-station Echo Canceller Test Equipment) in RMS is being considered. It seems likely that new functions will continue to be added to the RMS for some time to come.

References

1 Yayo, M.: *I/O Systems for AXE 10.* Ericsson Review *67* (1990):3, pp. 130–137.
2 CCITT Rec. M.605: *Routine Maintenance for International Public Telephony Circuits.* Volume IV, Fascicle IV.1.
3 CCITT Rec. G.703: *Physical/Electrical Characteristics of Hierarchical Digital Interfaces.* Volume III, Fascicle III.4.
4 CCITT Rec. G.821: *Error Performance of an International Digital Connection Forming Part of an Integrated Services Digital Network.* Volume III, Fascicle III.5.
5 CCITT Rec. O.11: *Maintenance Access Lines.* Volume IV, Fascicle IV.4.
6 CCITT Rec. O.22: *CCITT Automatic Transmission Measuring and Signalling Testing Equipment ATME No.2.* Volume IV, Fascicle IV.4.
7 CCITT Rec. O.27: *In-Station Echo Canceller Test Equipment.* Volume IV, Fascicle IV.4.
8 CCITT Rec. O.41: *Psophometer for use on Telephone-Type Circuits.* Volume IV, Fascicle IV.4.
9 CCITT Rec. O.152: *Error Performance Measuring Equipment for 64 kbit/s Paths.* Volume IV, Fascicle IV.4.

The DIAmuX System Series – Flexibility in the Access Network

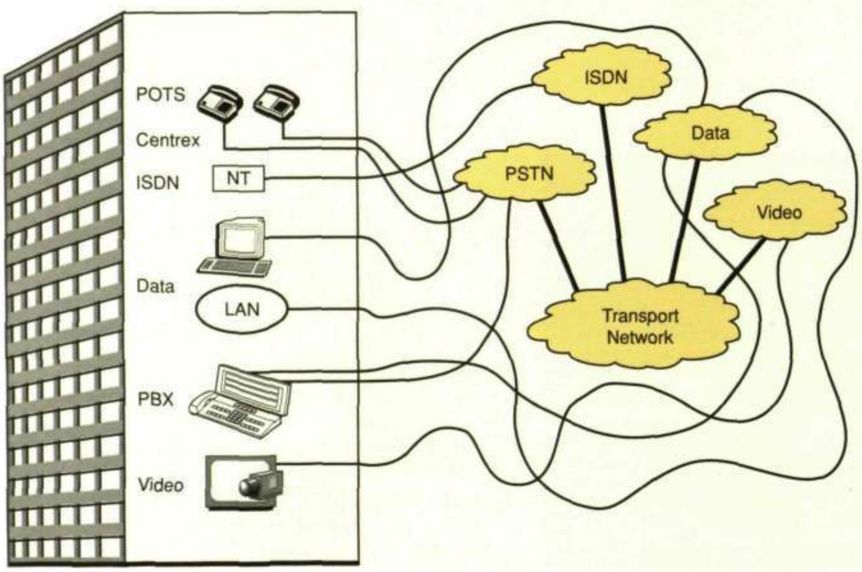
Niclas Frohm, Claus Lindholt Hansen and Daniel Madero

The telecommunications world is changing rapidly. Great efforts are put into the development of Managed Transport Networks like the Ericsson Transport Network Architecture (ETNA), in order to gradually replace today's fixed transmission systems. In the switching area, the concept of Intelligent Networks is gaining ground. When introducing Intelligent Networks systems, it is essential to have access systems for advanced customer services capable of following the migration path from PDH to SDH and from service-dedicated exchanges to multi-service nodes. That is why Ericsson now provides a series of access systems for use together with AXE exchanges, the ETNA transport network system and the TMOS telecom management platform.
The authors describe the DIAmuX System Series, a number of Flexible Access Systems for Integrated Access, developed to interwork directly with ETNA and AXE. DIAmuX is an Ericsson solution to managed and flexible, integrated access.

Access networks will be a prime area for investments in the 90s. Operators will have to offer their customers a wide variety of services over the access networks, today largely consisting of a dedicated copper pair for each individual service connection.

Business subscribers often require a mixed set of services from the operator: Plain Ordinary Telephone Services (POTS), data communication, leased lines and, perhaps, ISDN. Also, business subscribers often need to change the type of service required as well as the number of interfaces. Occasionally the customer requires that the services should be shared between physical locations or that all his business should be moved to some other place, always with a minimum of disturbances to the telecommunication services. A state of constant change is a proper description of the business world of today.

Fig. 1A
Dedicated transmission access for each service network makes it difficult to survey wiring in the access network



In the existing access networks, a request for some modification of the service provided usually requires physical changes in the network, which makes it difficult to handle the dynamic situation created by today's business. Ericsson's new flexible DIAmuX System, interworking with AXE and ETNA, provides a solution to this problem. DIAmuX supports features such as:
- integration of access to different services, using a common transmission channel
- effective management systems with options for partial customer control
- fast service provisioning
- digital exchange interfaces avoiding costly channel bank solutions
- simple installation and operation of remote locations in the access network.

Integrated Access:

Integrated access means that different services share the access network resources, which makes for cost-effective provisioning of a mixture of services, Fig. 1A and 1B.

As the demand for non-POTS services increases, and as faster service provisioning and improved service quality are requested by the customers, the access network must be capable of using the available bandwidth for the benefit of all services (interactive and non-interactive) in a cost-effective manner.

As an example, POTS, ISDN and data channels may be carried in the same 2 Mbit/s signals through the access and transport networks, separated at some point in the network in different directions towards the dedicated service nodes (e.g. exchanges for POTS/ISDN, cross-connects providing leased lines). The term used for this function is "grooming".

CLAUS LINDHOLT HANSEN
DIAX Telecommunications A/S
Denmark
NICLAS FROHM
DANIEL MADERO
Ericsson Telecom AB
Sweden

Part of the grooming function may be handled by the exchange itself. Protocols like British Telecom's DASS 2 and ETSI's V_5 handle both POTS and ISDN traffic.

Furthermore, integrated access can be seen as the alternative to private business networks implemented on leased lines supplied by the operator.

Where to use DIAmuX

The advantage of deploying DIAmuXes as Flexible Access Systems lies on the variety of services to be offered, the dynamics of customer requirements, network management requirements and other factors, Fig. 2. A non-expanding residential area where only POTS are demanded is no place to install a DIAmuX. However, if one or more of the facilities listed below are requested, the DIAmuX System Series might well be the solution:

- mixed services, with signalling
- fast provisioning and modification of services
- cross-connect and grooming facilities
- a wide range of site sizes
- outdoor installation
- a coherent management solution with a TMN interface
- fibre in both the primary and secondary access network.

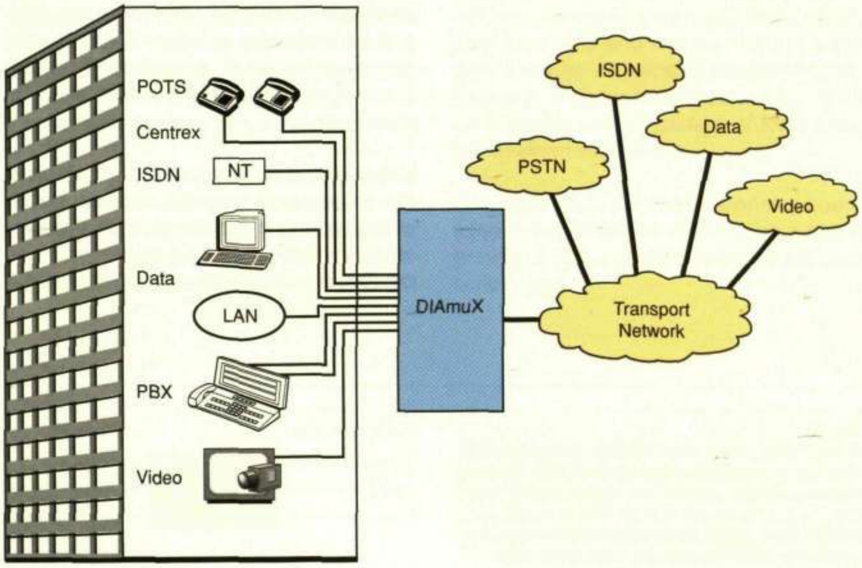


Fig. 1B
Integrated access provided by a Flexible Access System permits several services to share a common bandwidth

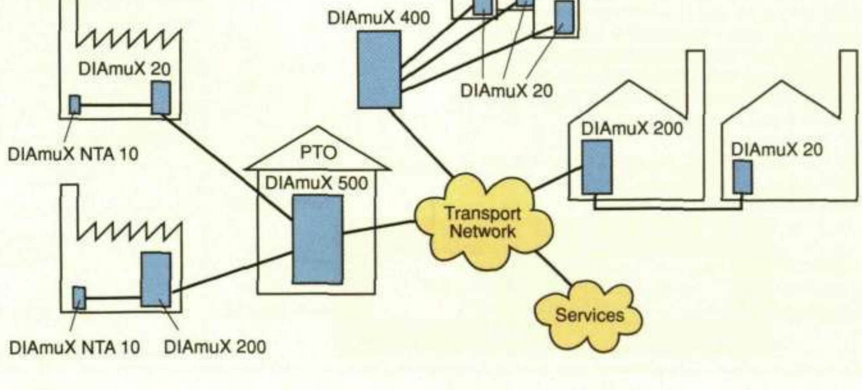


Fig. 2
DIAmuX can handle access to several services in large and small access node configurations with local and central management. Installations can be made on the operator's or customer's premises, or outdoors in roadside cabinets

The DIAmuX System Series

The DIAmuX System Series is Ericsson's solution to meeting the described service demands. The System Series consists of:
- DIAmuX 500 expandable from two to five 19" standard subracks
- DIAmuX 200 consists of one or two 19" standard subracks
- DIAmuX 400 a DIAmuX in a roadside cabinet. It contains two to four 19" subracks
- DIAmuX 20 a small DIAmuX for wall, shelf, desk-top or 19" rack mounting
- DIAmuX NTA 10 the very small "Network Terminal Adapter" for up to eight data interfaces.

DIAmuX 20 and DIAmuX NTA 10 are designed as remote units connected to, and controlled by, a larger DIAmuX. The hardware has been designed in order for small size and simple installation to be achieved.

DIAmuX 200, 400 and 500 are built up of the same hardware modules, which simplifies storing of spare parts and equipment for extensions. All of them can be equipped with either a large or a small circuit switch. The smaller module has 24 2 Mbit/s ports, corresponding to 768 switchable 64 kbit/s channels; the larger one has 96 2 Mbit/s ports and, consequently, provides 3072 64 kbit/s channels. The two switch boards are pin-compatible which makes it possible to start with a small and cost-effective configuration that can be upgraded with a larger switch when required.

All DIAmuXes can be deployed remotely in the network and managed from a central location. For this purpose, embedded signalling channels are used in the transport network to and between the DIAmuXes. The concept makes it possible to control several DIAmuXes from a workstation connected to one of the DIAmuXes. This workstation can also be used as a local O&M access.

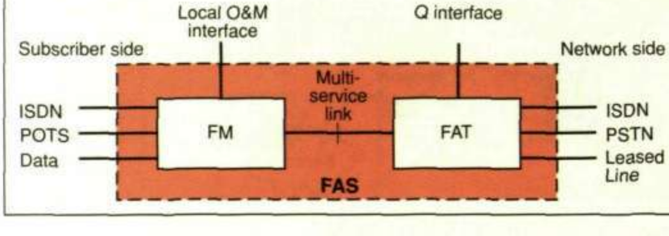
All DIAmuXes comply with the new EEC standards for EMC: EN 50081-1 and 50082-1.

DIAmuX 500

DIAmuX 500 is intended to serve as a large access node, e.g. for Fibre To The Build-

Fig. A1
A functional model of a Flexible Access System. The figure depicts a single-system FAS handling both subscriber access and service partitioning, grooming. The single-system FAS is a useful approach to public access points covering several subscribers or a single large subscriber

Fig. A2
FAS consisting of Flexible Access Termination performing the grooming function, while Flexible Multiplexers are used for the access function. This is a suitable configuration for fibre applications like Fibre To The Office or whenever a number of remote small access systems are used for multi-service traffic access, which is groomed in the FAT



Box1

Flexible Access System

The European Telecommunications Standards Institute (ETSI) is working on a standard for Flexible Access Systems (FAS) intended as a tool to provide for vendor independent integrated access. To do this, FAS supports – on the tributary side – a number of different physical interfaces and – on the aggregate side – various physical interfaces, Fig. A1 and A2. In order to obtain a direct (digital) connection to the public exchange it is crucial that also the signalling functions are handled.

A "core" function of FAS is the flexible mapping of the traffic channels and their associated signalling between the tributary and aggregate sides.

An FAS performs the access and the grooming functions. One or several Flexible Multiplexers (FM) provides access towards the subscribers, and a Flexible Access Termination (FAT) terminates the access network and performs the grooming function. These functions can also be brought together in one installation.

Ericsson is participating in ETSI's FAS activities in order to secure that the DIAmuXes will comply with the emerging standard.

Fig. 3A, left
The picture shows two DIAmuX 500 systems in two cabinets side by side. The one to the right is almost fully equipped with interface boards and with black dummy units placed in vacant board positions. The cabinet to the left is under installation. Note that there is no front cabling. All of the electronics is mounted on a swing-out frame, which provides access to the cabling from the front side of the cabinet. The cabinets have outlets for cables at top and bottom

Fig. 3B, right
The same two system as shown i Fig. 3A with closed cabinet doors. These two cabinets can handle around 800-900 subscriber channels depending on the type (ISDN, POTS or data)

ing (FTTB) or similar applications and also as a Flexible Access Termination (FAT) located at an exchange and grooming the services from remote DIAmuXes in the network.

DIAmuX 500 is built with 265 mm (6 HU) subracks implemented in 19" standard, Fig. 3. The printed board assemblies are in double Euro-size (16·23 cm). Hardware modules are designed with low building height - only 15 mm (3 TU). This means that the system can accommodate 28 interface modules in one subrack.

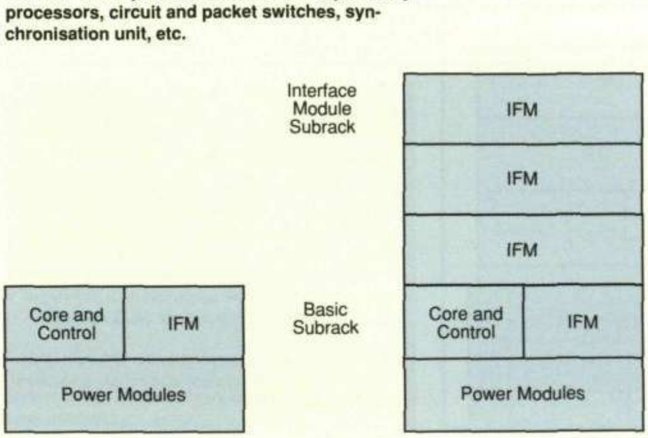
The maximum capacity is 92 printed board assemblies for 2 Mbit/s interfaces or 47 double-board groups for 14 subscriber lines each, giving a total of 658 subscriber ports. Present configurations need both subscriber channels and aggregate interfaces, which reduces the number of available subscriber connections. Example POTS: 448 POTS, 15 aggregate 2 Mbit/s.

DIAmuX 200

DIAmuX 200, Fig. 4, is designed to provide a cost-effective solution for smaller sites retaining all the functions of DIAmuX 500. The same SW and HW units are used in both systems. A typical application of DIAmuX 200 is as an integrated access node for a medium-size business or for a number of small businesses.

The upper capacity limit is 36 printed board assemblies for 2 Mbit/s or 18 double-board groups of 14 subscribers, equal to 252 subscriber ports.

Fig. B
DIAmuX 500 layout: Maximum configuration to the left, and minimum configuration to the right. The parts marked "IFM" are used for any type of *aggregate or tributary interface board*. "Core and Control" is the part that contains the supervisory processors, circuit and packet switches, synchronisation unit, etc.



Box 2

DIAmuX 500

DIAmuX 500 consists of three different subracks:
- the basic subrack containing core functions with *additional space for a suitable mixture of* Interface Modules up to ten 2 Mbit/s terminations or 70 ISDN basic accesses (or 64 kbit/s terminations)
- the Interface Module subrack (IFM subrack) with space for 14 master boards and 14 slave boards allowing connections of up to 196 ISDN basic accesses (or 64 kbit/s interfaces) or 28 2 Mbit/s terminations
- the power supply subrack accommodating power supply units and ringing generator. The DC/DC converters work in an n + 1 load-sharing fashion. A subscriber voltage booster, when *needed, is also mounted in the subrack.*

The minimum DIAmuX 500 configuration is one basic subrack and one power supply subrack, which can be expanded with up to three IFM subracks, Fig. B. *In the interface section in the basic* subrack and in the IFM subracks, any type of Interface Module can be placed in any position. The system is expanded by installing additional boards and subracks as capacity requirements increase.

Fig. 4A, left
A two-shelf *DIAmuX 200 system under* installation. As in the case of the DIAmuX 500 system, all electronics is mounted on a swing-out frame allowing access to the cabling from the front side. Access to the cabinet cabling is through top or bottom

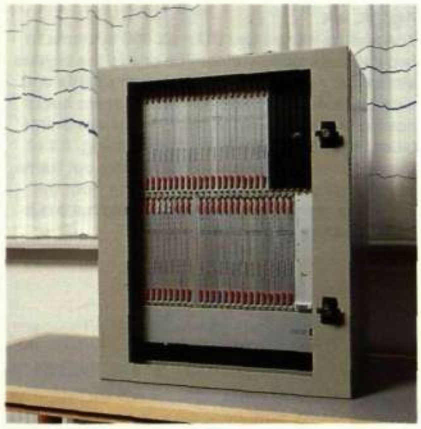


Fig. 4B, right
Installed DIAmuX 200 system capable of serving about 200 subscriber channels with a corresponding number of aggregate interfaces

DIAmuX 400
DIAmuX 400, Fig. 5, is contained in an outdoor cabinet with built-in climatic control. The MDF and electronics are mounted in an inner twin-cabinet with vertical separation and separate doors.

The cabinet has room for 66 boards for 2 Mbit/s or 33 double-board groups for 14 subscribers, equal to 462 subscriber ports.

DIAmuX 20
DIAmuX 20 is a flexible multiplexer capable of handling the access part of the Flexible Access System as described by ETSI. It provides a handy solution to distributed applications where various interfaces are needed, e.g. Fibre To The Building or Fibre To The Office. DIAmuX 20 has an integrated 2 Mbit/s fibre-optic line terminal as an option to the G.703 electrical interface.

If a number of DIAmuX 20 access systems are used in a network, they can be connected to a larger DIAmuX that serve as a Flexible Access Termination, FAT, performing a "grooming" function. This means that if the 2 Mbit/s lines from the DIAmuX 20 access systems are only partly filled – perhaps with several types of traffic – the "grooming" DIAmuX separates the traffic towards the relevant service networks. The cross-connect function makes it possible to terminate all unused channels from each DIAmuX 20, cross-connecting only those channels which are used. No bandwidth or subscriber numbers are wasted in the transport network or in the exchange.

The cabinet has space for eight interface boards, of which at least one is used for a 2 Mbit/s aggregate connection. The remaining seven positions are available for user interfaces. Depending on the type of interface, there are one, two or more ports per board. As an example, two standard POTS or ISDN basic access interfaces occupy one board, permitting 14 subscriber circuits to be connected via a 2 Mbit/s aggregate circuit.

DIAmuX NTA 10
DIAmuX NTA 10 is a flexible baseband modem with several interface ports. The standard DIAmuX ISDN basic access interface board is used to give an ISDN U–interface towards the superordinate DIAmuX. The multiplexing of data streams into the two B-channels follows the CCITT V.110 and I.460 recommendations. Up to eight user ports may share the capacity of the two 64 kbit/s B-channels. User port data bit rates ranges from 600 bit/s up to 64 kbit/s.

Like its larger family members, the NTA 10 is modular in design. Different user inter-

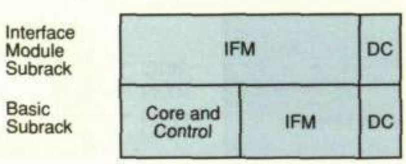


Fig. C
Layout of DIAmuX 200. The common section used alone is the DIAmuX 200 minimum configuration. Additional aggregate and tributary interface boards are placed in IMF

Box 3

DIAmuX 200

DIAmuX 200 has all the functions of DIAmuX 500 but the switch is not duplicated. DIAmuX 200, 400 and 500 can all be equipped with a large or a small circuit switch. This makes it possible to start with a small and cost-effective configuration which, when needed, can be upgraded with the larger pin-compatible switch board for 3072 channels.

The IFM subrack is identical to the DIAmuX 500, Fig. C, but the basic subrack also contains all necessary power units.

Fig. 5A, left
DIAmuX 20 with and without battery module. The electronics compartment is at the top and below it is the power module. The bottom part of the larger model is the battery module

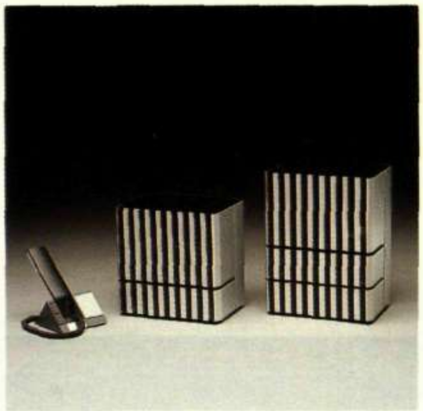
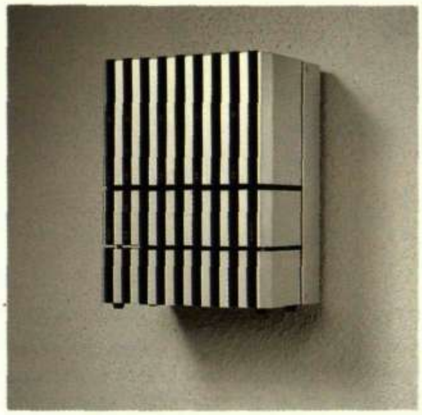


Fig. 5B, right
Wall-mounted DIAmuX 20. Dark transparent panels covers the LEDs that indicate status information. White text on panels mounted on each circuit board identifies the type of board and the indication given by each LED on the board. The front part of the cabinet can be tilted forward to provide access to plugs and cables in the back. The back of the cabinet also provides space for a fibre organiser



faces are implemented on interchangeable boards to obtain flexibility. The mechanical construction makes it look like a smaller and slimmer version of DIAmuX 20, Fig. 6.

The idea behind the NTA 10 is that the data interface should be available on the customer's premises as close to the data terminal equipment as possible. A DIAmuX 200, 400 or 500 may be placed at a public exchange location or on a customer's premises. In both cases the distance between the DIAmuX and the data equipment will normally be much greater than the 15-metre range of a V.24 interface. This may be the case even with DIAmuX 20. In order to overcome this problem, the data interfaces are provided at the NTA 10, which is then connected to a superordinate DIAmuX over a normal subscriber copper pair. In this way, copper already installed can be used for high-capacity multiplexed data.

Box 5

DIAmuX 20

DIAmuX 20 consists of a mother board, a processor board and up to eight interface boards placed in a case designed for wall-mounting or "desk-top" use. Brackets allowing 19" mounting are also available.

All services, handled by DIAmuX 200/400/500 are also provided on DIAmuX 20, managed through a 64 kbit channel as a remote unit to DIAmuX 200/400/500.

Downloading of software as well as configuration data is executed through the 64 kbit/s O&M channel in the 2 Mbit/s aggregate interface or through a V.24 local interface. The V.24 interface can also be used for certain O&M operations.

A separate power module is used to avoid mains voltage in the equipment case, supplying all DC voltages needed

Box 4

DIAmuX 400

DIAmuX 400, Fig. D, is housed in an outdoor cabinet with two inner cabinets, one for the electronics and the other for the main distribution frame, batteries and the mains rectifier. DIAmux 400 is built up of the same functional units as DIAmux 500 but only two IFM subracks can be used, which reduces the maximum capacity.

DIAmuX 400 is mains-operated. Batteries dimensioned for 6 hours of emergency operation in case of power failure are provided. This time may be extended if the DIAmuX is only partly equipped

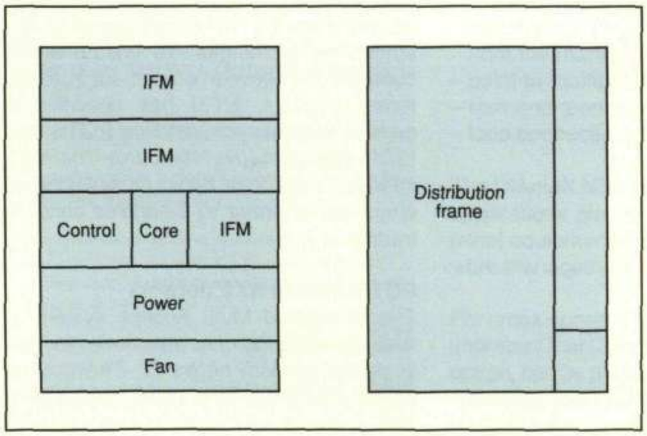


Fig. D
Layout of DIAmuX 400, outdoor cabinet with two inner cabinets

Fig. 6
DIAmuX NTA 10 is a small baseband modem for up to eight user interfaces. Mechanically it resembles DIAmuX 20, with dark glass panels covering the LEDs. The lower part of the cabinet covers cables and connectors, providing ease of installation, whether on a wall, a shelf or a table

Providing services with DIAmuX

In general, DIAmuX will cover all the services that an operator offers his customers. The variety of available interfaces for the DIAmuX family will continue to grow in phase with the increasing demands for services. And again, it should be noted that any new interface will easily be introduced in DIAmuXes already in operation, thanks to the flexibility of the system.

Configuration and provision of services can be made rapidly and efficiently if the access network elements have been given some excess capacity. If so, the operator can configure and activate interfaces using the management system without any need to visit the site where the equipment is installed. This requires, of course, that the customer himself can makes the physical connections or that it has been done in advance.

A key factor in providing services with access systems is the connection to the controlling exchanges. To ensure better compatibility between equipment from different vendors, ETSI has specified a generic interface for handling POTS and ISDN traffic, the V_5 standard. The $V_5.1$ interface describes basic multiplexing of channels, whereas $V_5.2$ handles concentration.

POTS with an AXE network
The Integrated Multi Access system in AXE 10 is a family of access nodes for use in public telecom networks. The access nodes offer a choice of radio, optical fibre and conventional copper technologies, enabling appropriate solutions to be selected for virtually any subscriber access network project. DIAmuX is one of these access nodes for the Integrated Multi Access system in AXE 10.

The DIAmuX is positioned as a node for integrated access to POTS, ISDN and data services (e.g. Frame Relay) with variable bandwidth up to 2 Mbit/s.

Plain Ordinary Telephony Services, POTS, are handled by DIAmuX 200/500/400 connected to a host AXE exchange via a 2 Mbit/s ESM interface. The signalling protocol used is an Ericsson proprietary protocol for Channel-Associated Signalling, RSM/ESM/CAS.

The RSM/ESM/CAS interface might be upgraded according to the ETSI $V_5.1$ and $V_5.2$ standard specifications when these are completed.

In the ETSI Technical Specification for V_5, the responsibility for maintenance of the subscriber line is shifted from the local exchange to a TMN system via a Q-interface. Therefore, ETSI V_5 does not include such functions. Since the definitions and standards of Q-interfaces seem to be remoter than the V_5 interface, Ericsson will offer a value-added V_5 interface with optional operation and maintenance functions.

POTS with non-AXE exchanges
The DIAmuX system concept is well suited for implementation of new signalling schemes. Most functions and protocols can be added through software updates in the main processor part or in the interface modules. For example, the same 2 Mbit/s interface board will handle ISDN LAP-D, RSM and ESM protocols – the protocol to be applied is activated by the management system.

In general, access systems and exchanges from different vendors will not work together. The RSM/ESM CAS signalling interface is a proprietary Ericsson implementation which only works with AXE exchanges. This is a situation which ETSI aims to improve with the V_5 standards. Even after V_5 has been widely accepted, it is obvious that a large number of exchanges without this interface will continue to be in operation (e.g analogue exchanges).

To connect DIAmuXes to non-AXE exchanges that do not have the ETSI V_5 interface, several options exist:
- The traditional way using a back-to-back multiplexer with a reverse analogue exchange interface. This is also possible with the DIAmuX, either directly or by means of the Ericsson CSM channel bank.
- A software update of the DIAmuX CAS signalling scheme to suit other types of exchange.
- Using market-specific protocols when available. An example is the UK, where the DASS 2 protocol is generally used.

POTS/ISDN via V_5 or PRA

When the exchanges in the network support the V_5 standard, both POTS and ISDN can be supported over this generic interface. Then a true multi-vendor environment for access and switching may become a reality with basic call control and service provisioning following a uniform standard. However, in the less established field of operation and management functions, Ericsson will provide optional extensions to the V_5 standard supporting coherent management solutions.

Providing ISDN is not just a matter of providing a U-interface board. The important issue is to establish a connection from the subscriber to the exchange that controls the service. This can be done in three ways:
- A back-to-back method presenting each individual U-interface on a copper pair to the exchange. Although this solution is the simplest one, practicable with ordinary multiplexers, it is unfavourable in terms of cost, installation and management (it entails the cost of three Basic Rate interfaces, but only one is provided).
- An ISDN Primary Rate interface between the access network and the exchange, which is made possible through advanced equipment like the DIAmuXes. This is a good solution as long as ISDN is implemented as an overlay network with gateways to the PSTN.
- An ETSI V_5 interface. This is relevant when the ISDN and PSTN merge together; that is, when one exchange handles both ISDN and PSTN calls. The ETSI V_5 interface solution is what DIAmuX is designed for – handling different switched services in parallel with leased lines, using a common channel signalling protocol.

Data interfaces and leased lines

Since the DIAmuX 500, 400, 200 and even the DIAmuX 20 have a cross-connect in their core function, the flexibility for leased line applications is high. The cross-connect is a non-blocking single-chip (DIAmuX 20) or single-board solution. Various types of connection are possible:
- $n \times 64$ kbit/s connections, $N \in \{1..31\}$. time-slot sequence integrity is preserved.
- 2 Mbit/s connection. The 2 Mbit/s interface board is then configured to allow time slot 0 to pass to the switch without being terminated. This is software-configured from the management system.
- point-to-multipoint connections
- split-and-monitoring connections
- loop connections for test purposes, etc.

The DIAmuX NTA 10 is ideal for small-size applications placed close to the data terminal equipment but managed by the network management system.

For cross-connect types of application it is important that DIAmuX 500 and 400, as an option, can be provided with duplicated circuit and packet switches in the core functional block.

Box 6

DIAmuX NTA 10

DIAmuX NTA 10 is a baseband modem with a U-interface for remote connection of data terminals to DIAmux 200, 400 or 500. DIAmuX NTA 10 supports up to eight data channels for data rates between 0.6 and 128 kbit/s. As the name suggests, it may be seen as a Network Termination and Terminal Adapter built together.

The bit streams from the data interfaces are multiplexed into the two 64 kbit/s B-channels in the U-interface. Data rate adaptation and multiplex framing is in accordance with CCITT recommendations V.110 and I.460.

A wide range of different port set-ups is possible, the limit being that the aggregated transmission rates, plus the framing overhead, must not exceed the 128 kbit/s aggregated bit rate of the two B-channels.

DIAmuX NTA 10 is configured and controlled from DIAX OM using the Embedded Operations Channel (EOC) in the U-interface. Configuration data are stored in the controlling DIAmuX and downloaded to the DIAmuX NTA 10 when its power is turned on.

Common functions, such as power, controlling microprocessor, multiplexing function, etc, are implemented on a mother board. Data channel interfaces are small plug-in modules.

The DIAmuX NTA 10 is suitable for wall-mounting or desk-top use, with all LEDs and status indications visible in both cases. All connections are made at one end of the unit.

DIAmuX NTA 10 uses a normal mains adapter for power supply. An internal DC/DC converter provides the necessary voltages for the data interfaces.

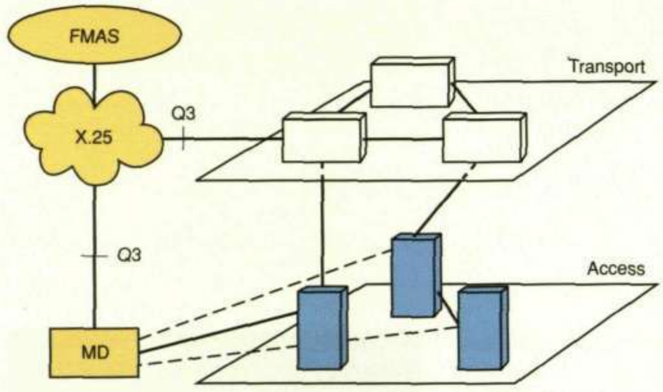


Fig. 7
Integrated operation and maintenance of access and transport networks. The figure illustrates how both ETNA transport network and access network elements can be managed by FMAS. The DIAmuXes are controlled via a mediation device (MD). DIAX OM mediation devices can be duplicated for availability reasons. DIAX OM may be used for certain O&M functions even when serving as a mediation device

Operation and maintenance

It is evident that efficient management is the key to an efficient access network. Some of the facilities provided by the DIAmuX family may be mentioned in this context:
- centralised control of access network elements
- quality control and measurements
- on-line alarm and event reporting
- fast service provisioning by management system messages
- efficient operator access security management
- flexible allocation of bandwidth.

An important advantage is that DIAmuX can be placed remotely and that sufficient management functions are implemented for this type of operation.

The DIAmuXes are self-configuring in the sense that the supervisory processor maintains an inventory of all the boards mounted, and of their status. This configuration is updated automatically when boards are mounted or removed. Services already active on a board will remain activated if the board is replaced in case of a fault. Intervention from the management system is not required, since the configuration and set-up of all boards are

Box 7
System concept

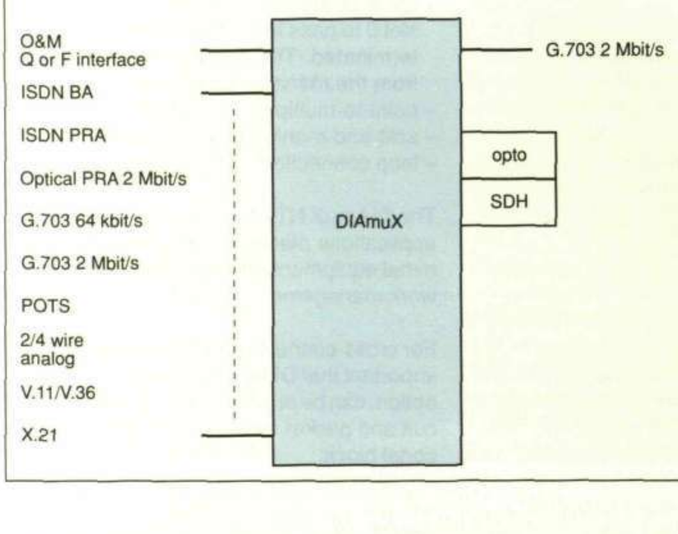
The system architecture of the DIAmuXes combines switching system features with a cross-connect function. This combination provides high system flexibility. It allows the connection of any input to any output as in a cross-connect system, but it also provides for the possibility of handling any type of protocol, be it channel-associated like the ESM/RSM signalling system or common channel as in ISDN Primary Rate Access or ETSI V_5. The DIAmuX System Series meets the requirements of all relevant CCITT and ETSI standards.

The signalling capability permits the connection of services directly to the switches over 2 Mbit/s interfaces. In this way the back-to-back or "channel bank" solution is avoided, reducing equipment cost and making operation and maintenance more efficient, Fig. E.

DIAmuX 500, 200, 400 and 20 all incorporate a circuit switch and the capability for packet-type signalling between the controlling processor(s) and the interface circuits. DIAmuX NTA 10 operates as a multiplexer and has no switch.

Main system features are:
- Circuit switching through a switch module with a capacity of up to 96 PCM systems equal to 3072 64 kbit/s channels. A few of these are used internally in the DIAmuX.
- A special unit in the core functional block handles packet signalling internally as well as externally to the system, using an embedded channel in the 2 Mbit/s highways towards each interface group module. This capability is used for internal operation purposes and for conveying network signalling.
- Signalling terminals are in most cases integrated on the 2 Mbit/s line termination boards. This means that a 2 Mbit/s interface unit can be software-configured to perform various protocols. In many cases, this can be done by sending a message from the management system; alternatively, software is down-loaded to the board.
- Compact design. A 19" subrack for interface modules accommodates 196 subscriber interfaces or 28 2 Mbit/s interfaces.
- Low power consumption. Maximum power is 250 W for DIAmuX 500 with 658 POTS connections. The actual consumption depends on the number and type of interface boards used, but the consumption for typical configurations is about half of the maximum wattage.

Fig. E
DIAmuX System Concept. On the aggregate side, several optical transmission alternatives are available from 2 and 8 Mbit/s fully integrated to higher-order PDH and SDH. "OPTO" and "SDH" blocks indicate that higher-order transmission equipment like STM1 or 34 Mbit/s PDH may be integrated with the DIAmuX installation



stored by each DIAmuX supervisory processor.

During operation, the configuration and status of DIAmuX boards are continuously monitored. Configuration data areas are checked with CRC at intervals of a few seconds.

A wide range of test and debugging facilities are available to localise hardware faults as well as software errors. The position and type of the faulty unit is immediately presented to the management system, Fig. 7.

DIAX OM: Element manager and TMN mediation device

The DIAmuX management system is called DIAX OM and works as an element manager and a mediation device providing a Q-interface. The Q-interface allows DIAmuXes in a network to be managed as network elements by a TMN system like Ericsson's Facility Management Application System, FMAS. The DIAX OM can be used as an element manager for several DIAmuXes in the network. Embedded channels in the access and transport networks are used for management communication.

With DIAX OM, the following functional areas are available:

- Configuration Management, e.g. handling hardware and software configuration of individual interface units as well as the core system. N·64 kbit cross-connect functions. Saving and dumping of configuration data.
- Fault management, e.g. on-line event monitoring and logging. Tracing and debugging facilities. Line measurements.
- Performance management, e.g. PCM line monitoring.
- Security management, e.g. creation of user groups with different access levels and passwords. Command logging.

DIAX OM can be used as the only O&M system in the access network, but a DIAX OM workstation can also be connected locally to a DIAmuX to provide partial local control over one or more DIAmuXes in the network. A DIAmuX can distribute alarms and events to more than one DIAX OM workstation in the network. This gives additional security for appropriate action if a customer's service fails, for example.

DIAX OM provides 16 relay outputs for driving external optical or acoustic station alarm devices. Configuration data and SW download program packages for a network of DIAmuXes may be stored on DIAX OM. The capability of having two or more DIAX OM systems connected to the network means that operations can be handled by personnel located at different sites.

FMAS

The Q-interface in the DIAX OM mediation device makes it possible to manage DIAmuX with the TMOS application system FMAS. The use of a TMN-based management system for the DIAmuXes in the access network is clearly advantageous.

FMAS manages the transport network, using an information model designed for network elements like SDH multiplexers and synchronous digital cross-connects.[1] Configuration, performance, fault and security management is handled on a network end-to-end basis by FMAS, extending the reach of the transport network through the management of DIAmuX.

Fig. 8
An example of an access network. A DIAMuX 500 is placed in the centre of a ring-type access network performing access and grooming functions. Remote access nodes are connected via PDH/SDH or, in the case of NTA 10, copper pairs. Access nodes can be installed in small sheds, in buildings or on the customer's premises. The access network provides the different user services over the common bandwidth supplied by the transport network

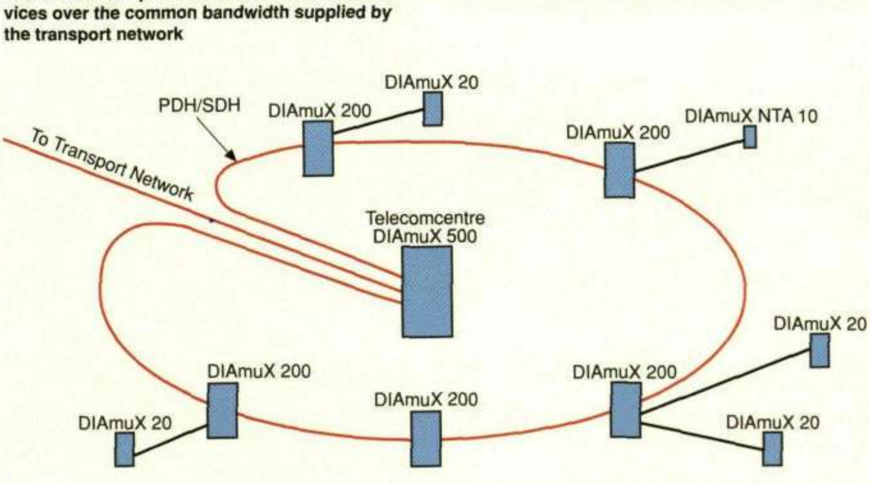
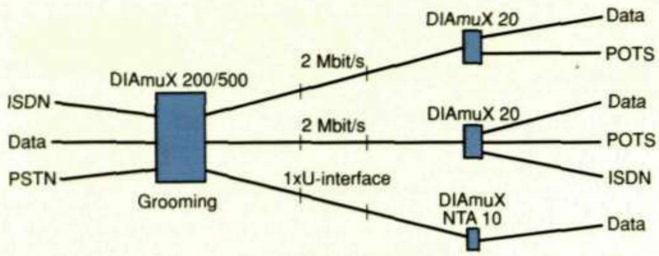


Fig. 9
An example of a typical access network consisting of Flexible Multiplexers (FM) and a Flexible Access Termination (FAT). DIAmuX 20 and DIAmuX NTA 10 are installed remotely with DIAmuX 200/500 at a central location. Exchange of interface boards in the DIAmuXes is the only hardware change required if the transmission system is changed from PCM on copper to HDSL or to fibre. The figure shows a network topology for Fibre To The Building, Fibre To The Curb, Fibre To The Home and similar applications



Bringing DIAmuX facilities into use in the TMOS/FMAS application will be through a phased introduction:

The first phase will see FMAS handling configuration, fault and performance management of the core functions in DIAmuX; that is, the cross-connect functionality. Additional access network functions will then be added – if appropriate in the form of a special application system for access network control. Examples of such functions are management of POTS, ISDN, sub-64 kbit/s services, and additional test and measurement functions.

What the DIAmuX can do for the network operator

With the DIAmuX System Series, Ericsson provides the operator with a uniform line of products for integrated access-products that are designed and suited to meet present and future needs in the access network. The concepts of service transparency, flexible access to the exchange, TMN management and the use of fibre in the access network are all provided for in the DIAmuXes.

There is a DIAmuX solution for any network topology and any node size. The cross-connect and signalling flexibility allows flexible network configurations. The NTA 10 provides a range from one to eight V.24 ports; the DIAmuX 500 will provide for several hundred POTS, ISDN and data connections. For sites of medium-size, DIAmuX 200 with one or two subracks, or DIAmuX 20, provides handy solutions.

The fact that the DIAmuX System Series covers the whole range of products from the same family means that hardware and spares are interchangeable, and that operation, maintenance and installation is similar for all components. Compatibility in all functions is guaranteed.

Furthermore, the close relation of DIAmuX to the AXE and ETNA systems ensures that coherent network solutions can be obtained both as regards switched services and management.

The flexibility of the DIAmuX concept simplifies updates of the access network with the emerging transmission technology.

Some examples of applications for access networks are given in Figs. 8, 9.

Summary

The DIAmuX System Series is Ericsson's solution to Integrated Access. It is based on flexible system design, well suited to meet the demanding requirements of a modern access network:
- integration of services with common bandwidth
- efficient management systems for fast provision of services
- generic and standardised exchange interfaces as well as Ericsson's proprietary solutions used today
- operation of systems in remote locations in the access network
- a wide range of services and site sizes
- multi-transmission technology.

The compatibility with ETNA and AXE systems ensures increased functionality in the complete network, both with respect to services and operation and maintenance.

References:

1 Tarle, H.:*An Operations Support System for Transport Networks.* Ericsson Review *67* (1990):4, pp. 163–182
2 Blume, J., Hägg, P. and Sundin, L.: *Control and Operation of SDH Network Elements.* Ericsson Review *69*(1992):3, pp. 62–77.

ERICSSON

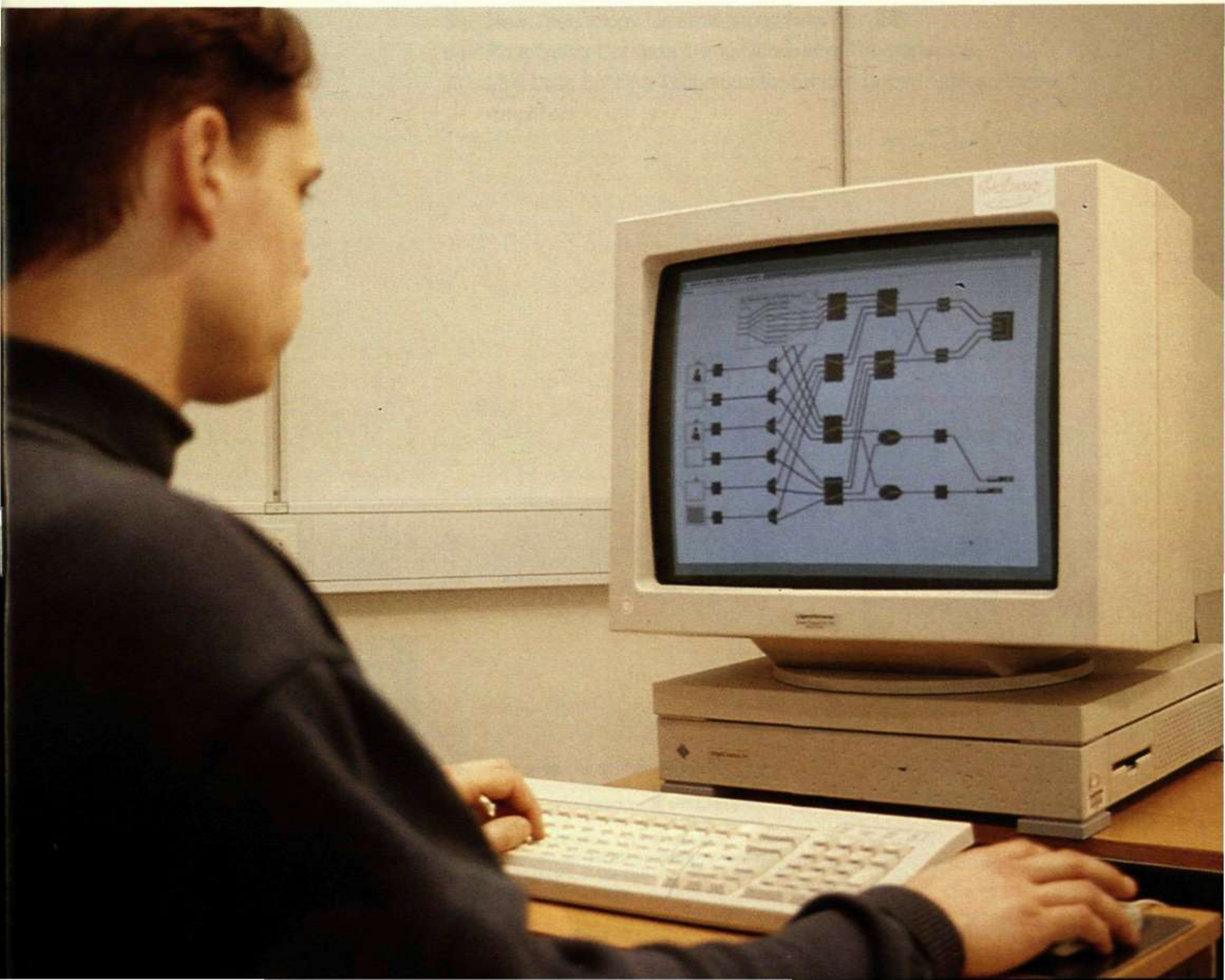
ISSN 0014-0171

Telefonaktiebolaget LM Ericsson

93156 Ljungföretagen, Örebro 1993

ERICSSON REVIEW

2 1993



ERICSSON REVIEW

Number 2 · 1993 · Volume 70

Responsible publisher *Bo Hedfors*

Editor *Per-Olof Thyselius*

Editorial staff *Martti Viitaniemi*

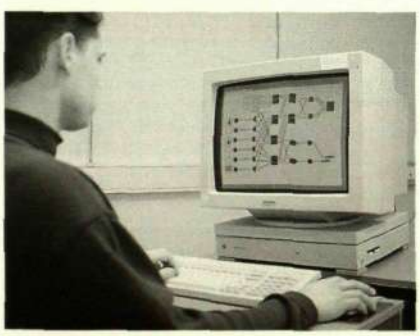
Subscription *Eva Karlstein*

Subscription one year $30

Address S-126 25 Stockholm, Sweden

Published in English with four issues per year

Contents



Cover
The RACE project Qscar – aiming to identify how and when optical switches shall be used in future broadband networks – is one example of an Erlang implementation. The picture shows the control workstation. On the screen an overwiew of the optical network with connected paths highlighted . In the subwindow, a view of one of the switch matrices.

RACE Open Services Architecture

Malcolm Key and Ulf Larsson

The specification and realisation of current telecommunication services tends to be intimately bound to a specific network architecture. Moreover, interactions between service modules are not always explicit, accessible or uniform, and tend to be optimised for a particular service. This has resulted in telecommunication networks and systems that cannot rapidly respond to changing customer requirements or exploit the advantages of new technology. The RACE Open Services Architecture (ROSA) project was established to address these problems.
The authors present an overview of the approach taken in the ROSA project.

RACE

RACE – R&D in Advanced Communication technologies in Europe – is an EC research program. The key concept in RACE is Integrated Broadband Communication – IBC – developed to denote the integration of broadband data, audio and image communication, expected to form the next generation of communications systems in Europe. RACE is based on collaboration among technical experts of the telecommunication sector – users, service providers, operators and vendors – in projects addressing different aspects of IBC.

The objective of the RACE Open Services Architecture (ROSA) project is the definition of an open architecture for Integrated Broadband Communications (IBC) services.

To understand the need for an open service architecture it is necessary to examine the current situation in telecommunication networks and services. Today's telecom networks are not flexible with respect to changes in service requirements or to the introduction of new services, and this is ascribable to a number of factors:
- the lack of a network-wide service architecture
- the limited capability for interworking between services
- the separation of service management from service control
- the close coupling between service control and network control
- the close coupling between service management and network management, and
- the complexity of interworking of networks based on different architectures.

All of these factors make service modification or introduction complex and expensive because of the redesign and reimplementation necessarily involved.

In view of both the standardisation bodies and relevant RACE functional specification projects, Integrated Broadband Communication (IBC) should support an extremely wide range of services. To support this view, a service-independent organisation of functions should be pursued as far as possible, to achieve the maximum flexibility with respect to service evolution during the lifetime of the Integrated Broadband Communication (IBC) system.

Within the above perspective, ROSA aims at providing an architecture for the specification, design and implementation of "open" services. Such an architecture must support the smooth introduction of new services, the graceful evolution of existing services, interworking between existing and new services, and increased independence of the services from network technologies.

The ROSA Approach

The ROSA project has defined an Open Services Architecture, a supporting methodology and an object model. Fig. 1 shows

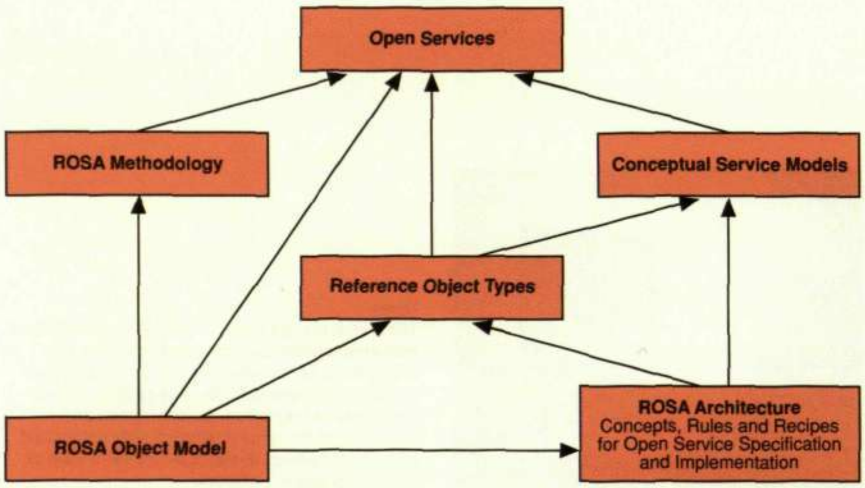


Fig. 1
The main components of ROSA are the ROSA Architecture, the ROSA Object Model and the ROSA Methodology. Reference Object Types are used as components in models of services, and Conceptual Service Models are recipes for modelling services

MALCOLM KEY
BT Laboratories
United Kingdom
ULF LARSSON
Ericsson Telecom AB
Sweden

the main components of ROSA; in this overview they are described through examples from a case study developed in the ROSA project. A short description of the ROSA Case Study is given in Box 1.

The ROSA notion of architecture aims at supporting an open service scenario. A key point in the ROSA architecture is its focus on telecom services, and only subsequently on network entities. This is consistent with our view that an open services architecture can only be stable if we concentrate first on what a system is meant to provide rather than on how a system will be internally organised.

The ROSA architecture is backed by a methodology. The ROSA Methodology guides a service designer in using the architecture to create "realisable" service specifications by stepwise refinement of usage-oriented service specifications.

The ROSA Object Model (ROOM) has been tailored to the specification of tele-

Box 1
The ROSA Case Study

A case study was undertaken in the ROSA project. The selected services in the ROSA Case Study were: a videogame service provided by a game centre, and broadband communication services supporting the use of multimedia services over a broadband network. The scenario used in the ROSA case study is inspired by game services currently available on the French Minitel terminals via the Teletel servers run by France Telecom. It makes the assumption that, by the time of the introduction of broadband communications, the following equipment will be available:
- A broadband user-network interface connecting the user's terminal equipment to the IBC network is available at the customer's premises
- A multimedia terminal with audio and video presentation capabilities.

The multimedia terminal must have the following features:
- a multi-window multimedia screen to display visual information including video
- stereo loud-speakers to reproduce aural game information
- a keyboard and a mouse or a joystick to control the game

Fig. 5 shows the organisations, or market players, involved in the Case Study services and their relationships. Entities in this *enterprise* model representing organisations are called enterprise *domains*. The three identified domains have different stakeholder roles in relation to each other. The actual services in the example are game services and IBC network services. These network services allow games to be played remotely. The organisations involved are:
- *The Game Player*, who uses the game service and uses IBC network services as a means of playing games. The game player receives charges or possible credits from game playing along with charges for other services provided by the IBC network services provider.
- *The Game Service Provider*, who delivers the game and is paid for it through the IBC Network Service Providers billing operations. For the network services used, the game service provider pays to the IBC network service provider.
- *The IBC Network Service Provider*, who provides network services that allow the game service to be delivered at the customers' premises and who is paid for these services by the Game Service Provider.

Fig. 6 shows a refined enterprise model used in the case study. Service requirements are expressed in customer-provider and user-provider relationships in this enterprise model. Examples of service requirements placed on the IBC Network service provider are:
- *Information Services*: Information on charging rules for game playing; before selecting a game service provider, the game player should be informed of the cost of using these services.
- *Service Selection Services*: Game service provider selection; it should be possible for the game player to select a game service provider.
- *Information Transport Services*: Data connections for interactive data communication and synchronised multimedia connections for video and audio streams.
- *Access Services*: The game player and the game centre need control and connection access to the IBC network.
- *Charging Services*: Charging of game players should be provided.
- *Security Services*: Authentication and screening of game players should be provided.
- *Service Management Services*: The IBC network service manager needs services to configure and deploy user services and to monitor Quality of Service.

Fig. 7 shows concepts used in the case study. These concepts are organised on the basis of aspects: Service Modelling, Access Modelling, Transport Modelling and Management Modelling from the ROSA architecture.

In the ROSA methodology, user view models and abstraction level models are used to deal with the complexity of telecommunication services. In the case study, the game service was modelled at three levels of abstraction, and separate user view models of the IBC Network services were constructed for the game player, game centre and IBC service manager. An integrated service specification based on the three views was then constructed.

Fig. 8 shows examples of objects used in the specification of IBC Network services in the case study. This specification contains a large set of object types.

Examples of these object types are:
- *User Agent*. The control access is modelled by the "user agent" concept from the ROSA architecture. A user agent handles user requests for network services. After authorisation, it provides the reference to the actual service. The user then has access to the services defined in the actual "user view specification". It also provides a reference to the actual logical terminal to be used.
- *Logical Terminal*. The transport access is modelled by the "logical terminal" concept from the ROSA architecture. Only multimedia transport aspects are modelled in the context of IBC network services. An instance of a logical terminal provides access points to connections.
- *Screening List*. A screening list will provide services that, given some information, provide further information about the user: credit rating, restrictions on services the user has access to, etc.
- *Service Directory*. In this scenario, the service directory contains information on game services - including charging rules - which the game player can access and use through the IBC network.
- *Call*. The call object type is an association of call parties that controls and manages connections. A call object also notifies a charging object of events that may have an effect on charging.
- *Control Call*. This is a specialisation of the object type call. The control call creates a Data Connection after negotiating between a calling party and a called party object.
- *Multimedia Call*. The multimedia call creates multimedia connections between two logical terminals after a negotiation between a calling party and a called party object.
- *Data Connection*. This object type is used for bidirectional transfer of data. It is part of the control call object

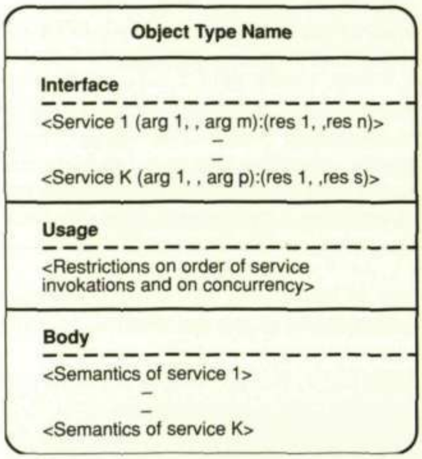


Fig. 2
An Object Type is a specification of an object using the ROSA Object Model (ROOM)

communication services, and it embodies concepts that make it particularly suitable for describing what a service does while abstracting away from the implementation-dependent details of how this service is provided.

The ROSA Object Model

Fig. 2 shows how a specification of an object class - an object type in the ROSA terminology - is structured. The ROSA Object Model (ROOM) is used both to describe services and to describe the components of the architecture in an object-oriented style. It has been designed so that telecommunication services can be described in an unambiguous and open way. The model consists of a set of concepts - relating to abstraction, specialisation, and composition - which are key features for constructing models. The main characteristics of ROOM are: *encapsulation, abstraction, composition/decomposition, inheritance,* and *reusability*. The main ROOM concepts are described in the following.

Object Type: An object type is a specification of the object properties (data types, services and behaviour). Object types consist of three parts: interface, usage and body. The interface describes what services the object offers to other objects. The usage part specifies restrictions on the possible order of invocation of services. The body describes the behaviour of the services.

Object: In ROSA, an object (object instance, or instance) is a model of a real-world phenomenon. An object offers services to other objects, and may also use the services offered by other objects. An object has a state which is a combination of internal data and the states of component objects, and which can only be changed by invoking services on the object. A service can be invoked by sending an appropriate message to the object offering the service. Objects are instances of object types and are created dynamically by a creator object.

Interface: This is the part of the object type specification which contains information needed in order to invoke the services of that type. It formally specifies service names, argument names, types and type of result. Multiple interaction patterns are allowed by combining notifications, service invocations and responses to service invocations.

Service: Objects offer services that can be invoked by other objects. The invoking object gives the service arguments; if appropriate, a result of the invocation is returned.

Usage: In ROOM, several services of a given object can be invoked concurrently. The usage part specifies restrictions on the order of invoking services and on internal concurrency.

Body: The body of the object specifies the behaviour (and thus the semantics) of the services offered by the object.

The ROOM promotes reuse of object types and objects through two main concepts: inheritance and composition.

In ROOM, object types can be created by the specialisation of more general object types through inheritance. Inheritance defines a relationship between types, and allows the definition of new types (subtypes) by adding new properties to those inherited from other types (supertypes). In this way a hierarchy of types can be built. Inheritance in ROOM can be both single and multiple.

In ROOM, object types may also be related through composition mechanisms. Two relationships are available in ROOM: *has part* expresses that an object contains instances of other types; *references* is used when reusing instances which are known by, but not a part of, the referencing instance.

Through these two relationships, the specification process appears as a combination of top-down or bottom-up approaches. This gives some flexibility for the definition of a design methodology. Besides, when associated with the modularity brought by encapsulation, these two combination mechanisms promote reuse, which is one of the openness requirements.

Service designers who use ROSA are required to specify a telecommunication service using the ROSA object model (ROOM). For this purpose, a set of Reference Object Types has been specified. These object types embody the core architecture and conform to the ROOM. Simple architectural concepts such as connections are expressed, logically, as single objects. More complex concepts, such as entire telecommunication services, are expressed as sets of interacting objects. Single objects are defined by single object types. Sets of interacting objects are defined by structured sets of object types. Architectural rules constrain the permissible interaction between instances of objects derived from these object types.

The ROSA Architecture

The purpose of an architectural framework is to appropriately structure an architecture, such that it is logically organised in a relevant way to the problem domain being addressed. Such a structure should position architectural components relevant to one another, guide the selection of appropriate components, and help to place boundaries upon an architecture. Two key principles provide the means of defining a framework: Viewpoints and Aspects. In ROSA, two viewpoints and six aspects have been identified.

Viewpoints are used to provide different levels of abstraction of the problem of interest. By concentrating on one viewpoint at a time, and ignoring the others, a designer can focus on a specific concern at a particular point in time.

Aspects, on the other hand, relate to a specific set of problems to be solved, or characteristics to be exhibited. In general, aspects pervade the different viewpoints.

The service designer's viewpoint is supported by the Service Specification Framework (SSF) and the system designer's viewpoint by the Resource Specification Framework (RSF). The service designer is concerned with specifying what a service provides, whereas the system designer is concerned with how a service can be provided to ensure that a design is realisable. Fig. 3 gives examples of concepts - object types - organised by aspects and viewpoints.

The Service Specification Framework (SSF)

The Service Specification Framework (SSF) provides concepts for the modelling of the "telecommunication-oriented" aspects of a service. By 'telecommunication-oriented' is meant the features found in everyday telecommunication services, such as calls, connections, charging, etc. A resultant specification using the SSF will prescribe what is to be offered to the users, in terms of how it is logically provided by interacting components. Distribution of components around a network is, by default, not a concern in the SSF, although requirements with respect to distribution may be expressed. The concepts in the SSF do not have any knowledge of the actual location of objects, which is to say that location and access transparency is provided.

Fig. 3
Aspects and viewpoints are used to structure the ROSA Architecture. Object types are organised on the basis of aspects and viewpoints

	Aspects				
Viewpoints	**Service Modelling**	**Access Modelling**	**Transport Modelling**	**Management Modelling**	**Object Support Modelling**
Service Specification Framework	Conference Session	User Agent	Video Connection	Alarm Log	Object Trader
Resource Specification Framework	Conference Bridge	Logical Terminal	Link	Basic Log	Object Creator

The Resource Specification Framework (RSF)

The Resource Specification Framework (RSF) provides concepts for the modelling of the "engineering-oriented" aspects of a service. By 'engineering-oriented' is meant the requirements for the implementation of a service in a network node. A resultant specification using the RSF will describe the requirements a service places on a system. The specification will not dictate how to implement, but will state what needs to be taken into account when implementing. What is required by a system to support the service is also expressed.

The primary concern when using the RSF is to identify which objects need to be implemented, and to model the distribution of objects in a network of interconnected nodes, in order to support requirements such as performance and reliability. The RSF will contain mechanisms to achieve some of the non-functional properties expressed by the SSF.

Aspects

The six aspects are called Service Modelling, Access Modelling, Transport Modelling, Management Modelling, Object Support Modelling, and Non-functional Modelling.

Access Modelling focuses on concepts needed to model service access. Access modelling provides concepts, such as user agent and logical terminal for modelling the behaviour required for access by each of the players associated with the service. Customisation of access is supported by the service profile and terminal profile concepts in the architecture.

Transport Modelling addresses concepts needed to model the control and characteristics of the information transfer between the users of the service. The architecture provides the notion of the trail and connection to model this conceptual area of the service. Through these concepts the quality of service parameters of the connection, the route and the behaviour of multimedia connections are controlled.

Service Modelling covers concepts that support the modelling of the core part of a service. The concepts of session, conferencing and directories are defined in the architecture to be used in modelling the service core functionality. The components used in this concept are mainly ser-

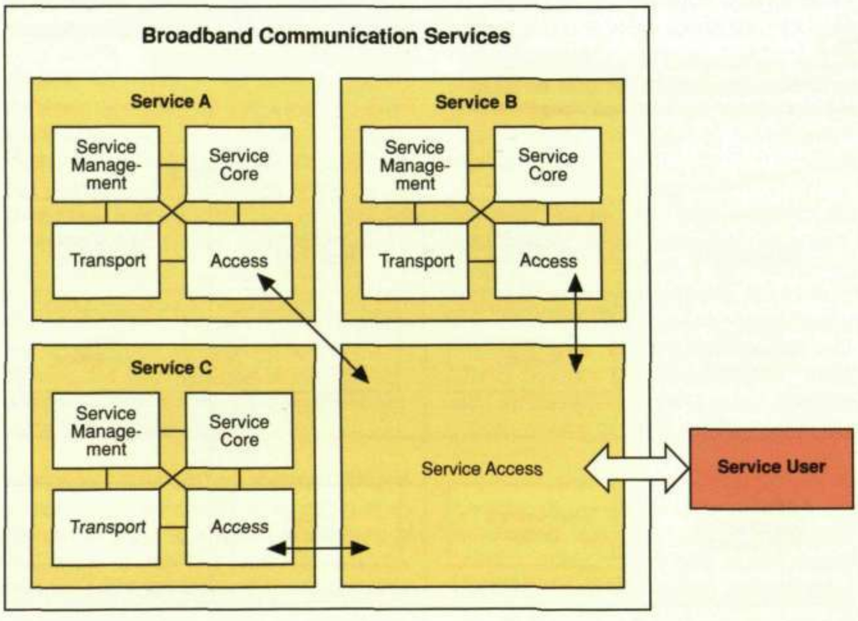


Fig. 4
Conceptual service models are used in order to fulfil the openness property of ROSA. A class of services will use a common conceptual model

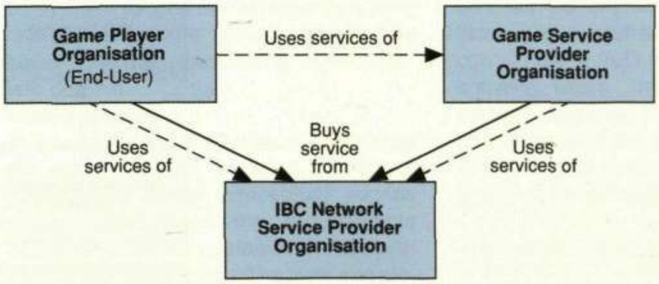


Fig. 5
An Enterprise Model shows the requirements customer and user organisations place on service provider organisations

vice-specific, since this is where all the unique capabilities of a service are defined.

Management Modelling. Each service needs a management component. Part of the management functionality is specific to the service; some parts are common to all services. Service-specific management includes the service's charging policies, service configuration and service performance analysis. Basic service management, on the other hand, involves the functionality required to monitor the general-purpose resources used by the service.

Object Support Modelling provides concepts for the reasoning about the dynamic properties of services. The dynamic properties of interest are creation and deletion of objects, and interaction between objects.

Non-functional Modelling provides concepts for the reasoning about availability, dependability and performance.

Conceptual Models
Conceptual models formalise the key concepts, separations and relationships that are crucial to open services and telecommunication systems. Conceptual models are used in the architecture to indicate how its basic concepts should be combined in the specification of services. An example of one type of conceptual model provided by the architecture is given in Fig. 4. The user in the figure can select services from a pool of broadband communication services provided to him in the IBC.

The ROSA architecture provides conceptual models to be used for classes of services and for classes of components used in services. Fig. 7 shows such a conceptual model. The ROSA methodology provides guidelines for how to use these architectural conceptual models to make conceptual models of specific services. A conceptual model of a specific service contains objects described by informal or formal specifications of object types. An example of such a model is shown in Fig. 8.

The ROSA Methodology

An architecture must be backed by a methodology; that is, a set of guidelines which assists the designer throughout the overall development process. The ROSA Methodology guides a service designer in using the ROSA architecture to create "realisable" service specifications by stepwise refinement of usage-oriented service specifications. It uses the ROSA Object Model in conceptual modelling and specification of services through several levels of abstraction, until the service is expressed in terms of objects that are realisable in an existing or perceived target network. The ROSA Methodology addresses three principal activities:
- Service Analysis
- Service Specification
- Service Implementation.

A short description of service analysis and service specification is given here.

Service Analysis
ROSA advocates the specification of a telecommunication service using the object-oriented approach at several levels of abstraction.

At the highest level of abstraction, a service is viewed in the context of an enterprise or a federation of enterprises. This abstraction helps the designer to formally capture the requirements of each player who is a stakeholder in the service.

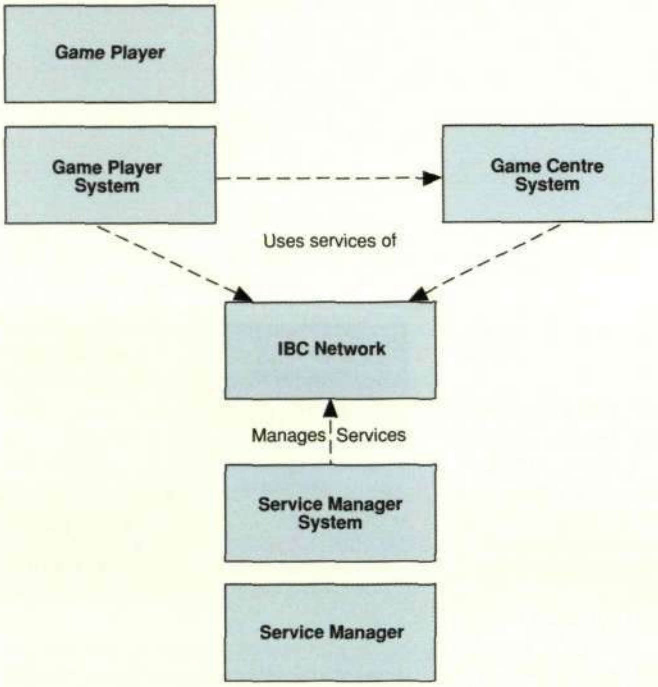


Fig. 6
The Enterprise Model is refined to show user and management services required from logical systems

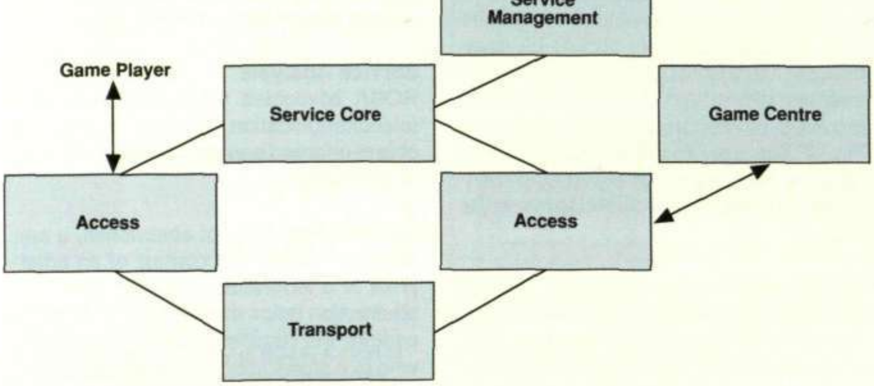
In this activity, an enterprise model that provides the context for the modelled service or services is constructed. The enterprise model comprises two levels of abstraction. The top level shows service customer and service provider organisations and their requirements and relationships. Stakeholder roles are used in identifying and structuring requirements. Actors in an enterprise model can assume three different service stakeholder roles: provider, customer and user. The enterprise model may express policies on the distribution of functions on enterprises or parts of enterprises. The concept of a domain as a means of modelling such a part of an enterprise model is provided by the ROSA architecture. Fig. 5 shows the enterprise domains identified in the ROSA case study.

The enterprise model is then refined by identifying users, service managers and logical systems as parts (sub-domains) of enterprise domains in the model. Users are agents for customer organisations, and service managers are agents for provider organisations. A logical system is a subdomain in the enterprise model owned by one provider or customer organisation. One or more services are now identified, each in the context of a logical system owned by a service provider. For each service, a conceptual model will be constructed. Fig. 6 shows the logical systems and users identified in the ROSA case study.

Rules and guidelines for logical partitioning of functions of a service are provided by the ROSA architecture. Fig. 7 shows the functional partitioning used in the ROSA case study. The ROSA architecture provides conceptual models for certain classes of services. The service designer selects a conceptual model for the actual class of service. Using the ROSA methodology a conceptual model for the actual service is constructed on the basis of the conceptual model from the architecture.

The first step is to build a local conceptual scheme of the actual service for each user including the service manager representing the provider and optionally the customer service manager(s). The Service Specification Framework of the ROSA architecture provides the service designer with some of the concepts to be

Fig. 7
Conceptual models for the different user services are constructed on the basis of models provided by the ROSA Architecture



used in the service model; other concepts are defined by the service designer by classification of phenomena in the problem domain.

In a second step, the local user schemes are integrated into a global conceptual model. The global conceptual service model incorporates the requirements which the user, the customer and the provider place on a service. The local and global conceptual models are constructed in the service analysis activity using object-oriented techniques.

Service Specification

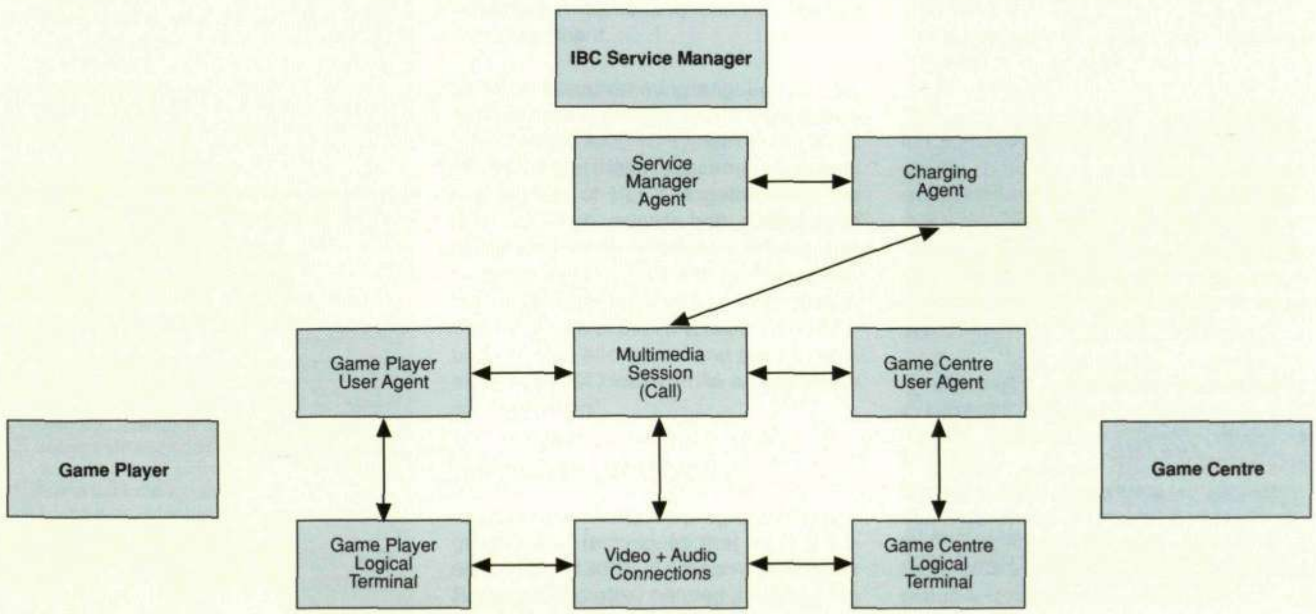
The global service model is formalised by specifying object types representing concepts in the service. A specification on this level of abstraction is called a service level specification in ROSA. The service designer uses component objects in order to specify the service in such detail that users, customers and providers can validate that their requirements are expressed in the service specification. Fig. 8 shows examples of such object types specified in the ROSA case study.

The service level specification is refined to take additional requirements from the enterprise context into consideration. A typical requirement is that "part services" provided by other service providers must be used. Conceptual modelling and specification of a part service is done in the context of a refined enterprise model showing users, managers and logical systems involved in the part service. A specification of the actual service refined to this level of abstraction is called a logical level specification.

The lowest level of abstraction considered in ROSA is when a logical system corresponds to a node or terminal in a network. A less abstract model of the target network provides the context for the service specification at this level of abstraction. The service designer now focuses on quality of service requirements and interconnectivity issues, using concepts from the Resource Specification Framework in the ROSA architecture to refine the specification into a less abstract realisation level specification.

A realisation level specification contains a set of object types which provide a basis for defining standard components and interfaces. If the ROSA architecture is applied when a telecommunication service is modelled, the objects and services they model will be open. It is the ability to create logical service models using architectural concepts coupled to object-orientation that gives ROSA its openness characteristics.

Fig. 8
Object Types are used to specify services according to the ROSA Architecture using the ROSA Methodology



Box 2

Acknowledgments

The ROSA project is partially funded from the European Community, and this paper is an edited summary of the collective work of many individuals who are participating in the project on behalf of the companies named below.

Architecture Projects Management Limited (UK), British Telecommunications plc (UK), CSELT (Italy), Ericsson Telecom AB (Sweden), Ericsson Telecom (UK) Ltd, FRANCE TELECOM, FUB (Italy), GMD FOKUS (Germany), IBM France, Norwegian Telecom, PTT Research (NL), Synergie Informatique et Développement (France) and TFL (Denmark).

Conclusions

The ROSA project has produced an architecture which is designed to be practical and compatible with major architectural initiatives and related RACE projects. The technical results of ROSA are documented in five deliverables (ROSA 1 – ROSA 5). The object-oriented architectural approach to the specification of generic service components is now being used in relevant RACE projects. ROSA spearheaded this change of view. ROSA was also one of the first projects in RACE to adopt the object-oriented architectural approach to distributed processing, used in ISO/ODP and originally defined by ANSA.

The ROSA project terminated, as planned, in December 1992, but a RACE II project, CASSIOPEIA, is now defining an Open Services Architecture, OSA, for Integrated Broadband Communication based on the ROSA results. This project has forged a relationship with the RACE Consensus Management project which will guide the OSA standards proposals to ETSI, CCITT, ISO and other standardisation bodies.

In parallel with ROSA, Bellcore has been defining an Information Networking Architecture (INA) with similar scope and objectives to those of ROSA. Indeed, the INA architecture has followed the ROSA approach in applying object orientation and open distributed processing techniques to telecommunication architecture.

ROSA Partners are now exploiting ROSA concepts and technology in a global telecommunication system architecture initiative called Telecommunication Information Networking Architecture (TINA). A consortium, TINA-C, has been formed by ROSA Partners, Bellcore, NTT and other organisations with the intention of defining an open services architecture with worldwide application.

References

1 ROSA 1:*Specifying IBC Services Using the ROSA Architecture,* RACE Deliverable 93/BTL/DNR/DS/A/003/b1, May 1992.
2 ROSA 2:ROSA *Architecture, Release Two,* RACE Deliverable 93/BTL/DNR /DS/A/005/b1, May 1992.
3 ROSA 3:*Foundation of Object-Orientation in ROSA,* Release Two, RACE Deliverable 93/BTL/DNR/DS/A/010, May 1992.
4 ROSA 4:*ROSA Case Study,* Race Deliverable 93/BTR/DNR/DS/A/011/b1, May 1992.
5 ROSA 5:*The ROSA Handbook,* Release Two, RACE Deliverable 93/BTL/DNS /DS/B/013/b1, September 1992.
6 RACE:*Research and Development in Advanced Communications Technologies in Europe,* RACE '92, Commission of the European Communities, March 1992.
7 ODP:Draft Recommendation X.903: *Basic Reference Model of Open Distributed Processing,* Part 3: Prescriptive Model, ISO/IEC JTC1/SC21/WG7 /N7055, June 1992.
8 ANSA:The ANSA *Reference Manual, Release 01.00,* APM Ltd., March 1989.
9 INA:Fleck, J.J., Liou, C.C., Natarajan, N. and Phillips, W.C.: *The INA Architecture: An Architecture for Information Networks,* Bellcore, TINA 92.

Erlang — A New Programming Language

Bjarne Däcker

BJARNE DÄCKER
Ellemtel Utvecklings AB

The telecommunications industry is highly dependent on computer software, the development and maintenance of which account for a very large part of system costs. It is therefore essential that the productivity in software design is improved by using the most powerful technology available.

Ericsson has developed the Erlang programming language, which is a declarative language that supports concurrency, distribution, error handling and code updating in running systems. Erlang has been used for research and prototype projects – including several RACE projects – for five years, and the extensive experience thus gained has proved its ability to drastically reduce system development efforts. Erlang will be available for commercial purposes in 1993.

The author describes the systematic prototyping activities which led to the design of the Erlang programming language. To illustrate the use of Erlang, two RACE projects – Oscar and Biped – are also described.

Programming telecommunication switching systems involves a unique combination of problems:
- Continuous operation for many years
- Handling a very large number of concurrent (parallel) activities
- Interaction with hardware
- Actions to be performed at a certain point of time or within a certain time
- Software maintenance (reconfiguration, etc) without stopping the system
- Stringent quality and reliability requirements
- Fault tolerance both to hardware failures and software errors
- Very large software systems
- Systems distributed over several computers
- Complex functionality such as feature management.

The PLEX programming language and the APZ computer architecture[14] were developed for the AXE 10 system in the early 1970s to meet these requirements. Blocks and signals of PLEX, together with the APZ hardware, provide both a method of structuring large software systems into modules and a mechanism for implementing massive light-weight concurrency. The hardware of the APZ allows reconfiguration, size alterations and the replacement of PLEX blocks while an exchange is in operation.

Language prototyping

In order to evaluate the programming languages and techniques that were available in the 1980s, Ericsson's Computer Science Laboratory created a prototyping environment consisting of a Line Interface Module, LIM, from the MD110 system. This LIM was modified so that it could be controlled by a VAX 11/750 running Berkeley UNIX.

During the period 1982-84 a series of experiments, based on this equipment, was performed. The experiments dealt with the programming of POTS (Plain Ordinary Telephony Service), using more than 20 different languages and paradigms, e.g.:
- Conventional "imperative languages" such as Ada and Concurrent Euclid
- Object-oriented languages such as CLU and a frames-based expert system
- Rule-based system
- Functional languages such as ML and Hope
- Logic programming languages such as Prolog.

This work was presented at a symposium in Eindhoven in 1986.[11] The shortest and clearest programs in the study were those written in functional and logic languages, the so-called "declarative languages". These languages were then largely research topics. They could not handle concurrency and were not suited for large systems.

Concurrency in modern operating systems or languages is handled by processes. The process (or task) concept in languages such as CHILL or Ada would be highly suitable for telecommunications applications too. However, in implementations of these languages the overhead for each process instance restricts their usability for massive concurrency.

Box A

Facts about Erlang

- Erlang is a programming language for implementing robust, fault-tolerant real-time systems.
- Erlang is available for many real-time operating systems and workstation operating systems.
- Erlang allows programs to be updated in a running system without stopping the system.
- Erlang has three error-detection mechanisms which simplify the design of robust and fault-tolerant systems.
- Erlang provides modules for structuring large programs.
- Erlang systems are dynamically linked at load time, and modules can be loaded individually.
- Sequential Erlang is a functional programming language with a pattern-matching syntax. This results in very short and succinct programs.
- Erlang programs are usually much smaller than equivalent programs written in conventional high-level languages
- Erlang has light-weight processes which provide an efficient means of programming concurrent systems.
- Erlang processes communicate through asynchronous message passing.
- Erlang makes for easy and efficient programming of distributed systems.
- Erlang has a foreign-language interface which allows programs written in other languages to be called from Erlang.
- Standard Erlang libraries include an ASN.1 interface, interactive graphics, mathematical subroutines and many other useful functions.
- Erlang is a small language which is taught in a four-day course.
- Erlang is available free of charge for research and educational purposes.
- Erlang is available on commercial terms for product development.

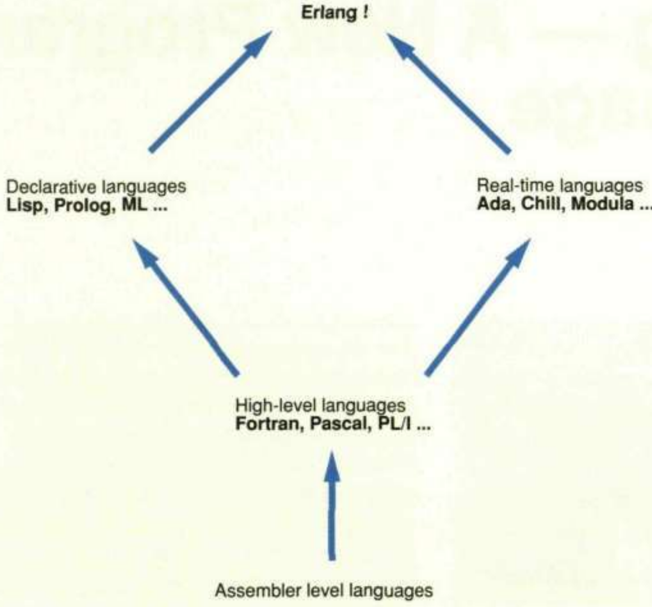


Fig. 1
Erlang merges two different programming language traditions: Declarative languages, which represent the development of traditional high-level languages, and real-time languages, which represent the specialisation of high-level languages for real-time control systems

None of the languages used in the experiments provided a good solution to the problem of updating software in running systems; nor did any of them provide a powerful means of programming error recovery.

The "ideal" language was thus a declarative language combined with light-weight concurrency. Experimentation continued by extending Prolog with a process concept. However, two aspects of Prolog make it unsuitable for real-time systems:
- Back-tracking – which allows a computation to be "undone" and retried to obtain the next result. This is difficult to use for hardware, i.e. a given order to hardware cannot easily be undone.
- Garbage collection – automatically reclaiming memory that is no longer used – stops the system periodically for a few hundred milliseconds. This is not acceptable for real-time systems. Schemes for incremental garbage collection are well known but are not widely available in Prolog implementations.

Experiments were made using logic languages which support concurrency, in particular the language Strand.[13] Strand, however, has the drawback of being too concurrent. All goals are executed in parallel and it becomes complicated to program actions which must be carried out in a specified order, for example when controlling hardware.

These experiments resulted in the design of a new programming language. Following the tradition of naming programming languages after distinguished mathematicians, it was called Erlang, Fig. 1.

The first implementation of Erlang was an interpreter written in Prolog. This facilitated language changes and experiments, but the implementation was too slow and could only be used for building prototypes.

Applications prototyping

In 1988, a team at Ericsson Business Communications, EBC, was working on a prototype of a new software architecture for PABX applications. The goals of the project were improved functional modularity, better ways of handling feature interaction and, most importantly, reduced design and documentation work. This team decided to base their prototype system on Erlang. Execution speed was a minor issue; the team was more concerned with the ease by which they could program and experiment with the software architecture.

The prototype which was the result of this work comprised about 35,000 lines of Erlang. The combination of a new software architecture and the new Erlang implementation technology resulted in a far smaller software system than that of existing systems of the same functionality. The design effort required was carefully monitored and compared with current techniques and was shown to be at least one order of magnitude less.[12]

During this project the language underwent several major revisions, and the user team showed great patience in putting up with often backwardly incompatible changes. The design of the Erlang programming language was thus performed jointly with the design of the first large application in Erlang. Comments from the user team directly influenced the design of the language, and the effects of changes in the language could be directly monitored by observing programs written by the users.

Defining Erlang

Erlang is a small language. Features which the user team did not use were removed, and features were added only when it was found that their absence forced the users to write complicated or unclear programs. Language constructs which were difficult to implement efficiently were excluded. Features such as higher-order functions, currying and "lazy evaluation", which are found in many functional languages, were excluded. Lightweight concurrency, primitives for error handling and mechanisms for loading code into running systems were included.

Code examples are given in Fig 2. Sequential Erlang is an untyped functional language. Concurrent processes are created by the *spawn* statement, and processes communicate through asynchronous messages passing in a way similar to that of SDL.

Links can be created between cooperating processes. If an error occurs in a process, it will be abnormally terminated. By default, all processes linked to an abnormally terminated process will also be terminated. This default action can be overridden so that "operating system" processes can monitor "application" processes and clean up and restart them. A well-structured system can be programmed so that it automatically recovers from errors in the application processes and only requires that vital operating system processes, e.g. resource managers, are absolutely correct.

In 1990, Erlang was recommended as the programming language for use in building prototypes within Ericsson. It was presented at the XIII International Switching Symposium, May 1990.[4]

A detailed description of Erlang is given in the language manual[5] and the standard textbook.[8]

Related languages

In 1980, Ericsson released the programming language EriPascal.[10] EriPascal is based on Pascal extended with concepts largely inspired by CHILL, such as modules and processes. EriPascal is widely used in many different projects in Ericsson and it has had a large influence on the design of concurrency in Erlang. For example, in EriPascal processes commu-

Fig. 2
The same problem solved in C (left) and Erlang (right). The Erlang program is much shorter than the C program. Erlang programs can be developed quicker than equivalent programs in conventional high-level languages

```
struct address { char *street; int number; };
struct person { struct address *addr; char *telno; };
char *check_malloc(size)
{
        char *s;
        if ((s = (char *) malloc(size)) ==0) exit (1);
        return s;
}

struct address *make_address(street, num)
char *street;
{
        struct address *a;
        a = (struct adress *) check_malloc(sizeof (struct address));
        a->street = (char *) check_malloc(strlen(street) +1);
        strcpy(a->street, street); a->number = num;
        return a;
}

struct person *make_person(addr, tel)
struct address *addr; char *tel;
{
        struct person *p;
        p = (struct person *) check_malloc(sizeof (struct person));
        p->telno = (char *) check_mallock(strlen(tel) +1);
        p->addr = addr; strcpy(p->telno, tel);
        return p;

}
do_something()
{
        struct person *p;
        p = make_person(make_address("Big street", 23), "124679");
}
```

```
X =     {person, {address, "Big street",23),
                 {telno, [1,2,4,6,7,9]}}.
```

nicate through asynchronous message passing, and EriPascal also has an analogous notion of linked (connected) processes.

Programming languages such as C or C++ are commonly used for implementing real-time systems. However, C and C++ are sequential languages which require either operating system support or language extensions to handle concurrency. An Erlang system running on UNIX, for example, can easily cooperate with programs written in these languages.

Further developments

Early implementations of Erlang were interpreters written in Prolog. As has been mentioned already, the execution speed and the fact that Prolog implementations pause to do garbage collection restricted the applicability of these implementations for large-scale usage.

Faster implementations were made by using compilation techniques similar to those used for languages such as Prolog, Fig. 3. The lack of back-tracking makes Erlang easier to implement than Prolog, but the implementation of light-weight concurrency and distribution poses other problems. Different techniques have been used to realise incremental real-time garbage collection.[6,7]

The impetus behind the design of Erlang was the programming of telecommunications software. However, the combination of declarative programming and light-weight concurrency has turned out to be very useful for programming a much wider range of applications.

Thus, the compiler for Erlang is written in Erlang itself, whereas the kernel has been written in C. Several tools have been written for Erlang programming – for example, an ASN.1[16] compiler which generates Erlang code.[17]

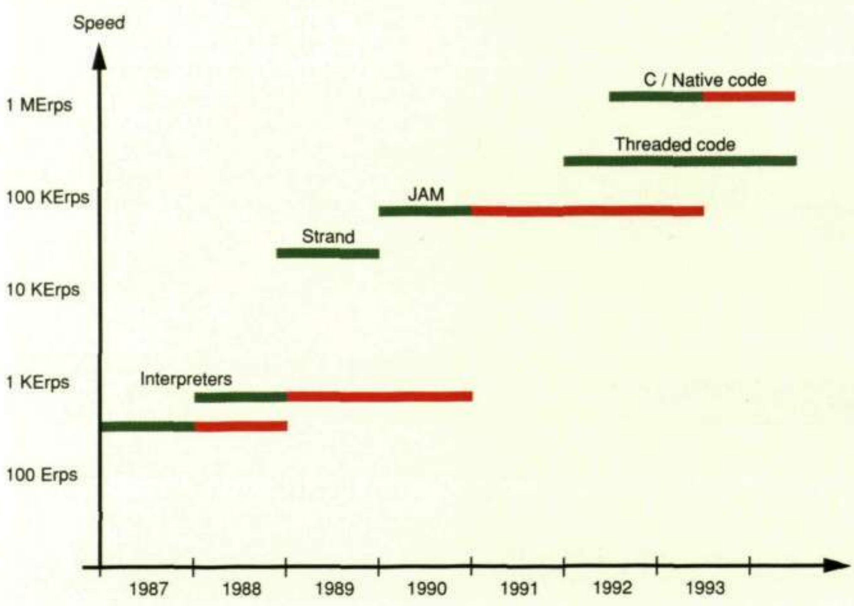
An unexpected application domain is interactive graphics for creating user-friendly interfaces. Library functions have been written to interface Erlang to the X Window system directly and using Interviews. Windows on the screen are coupled to Erlang processes so that inputs to the system are received in Erlang as ordinary messages, and output to the X Window system is sent as messages. The graphics user interface was used as part of the support system for Erlang. Graphics has opened up other application domains and increased the usefulness of Erlang for prototyping. A large prototype was the development at TELI AB of a demonstrator for a new paging system. Another prototype is a system for controlling a cordless switch, described in a parallel article.[1]

Communication between Erlang systems (and other systems) has long been possible by means of several methods; for example, through the use of sockets when Erlang is run on UNIX. Erlang was designed to be used for programming distributed applications, and distribution has now been integrated into Erlang implementations. This means that communication between processes, linking of processes, error handling etc can be done transparently between Erlang systems running on different computers, and even on different types of computers.

RACE Oscar and Biped

Erlang has been made freely available for research and education, to gain feedback from a larger user community. It has been used in several RACE projects. RACE

Fig. 3
Performance of the Erlang system. The green lines represent experiments, the red lines production systems. "Erps" means Erlang reductions per second. The fastest Erlang system, shown at the top of the diagram, runs as fast as unoptimised C



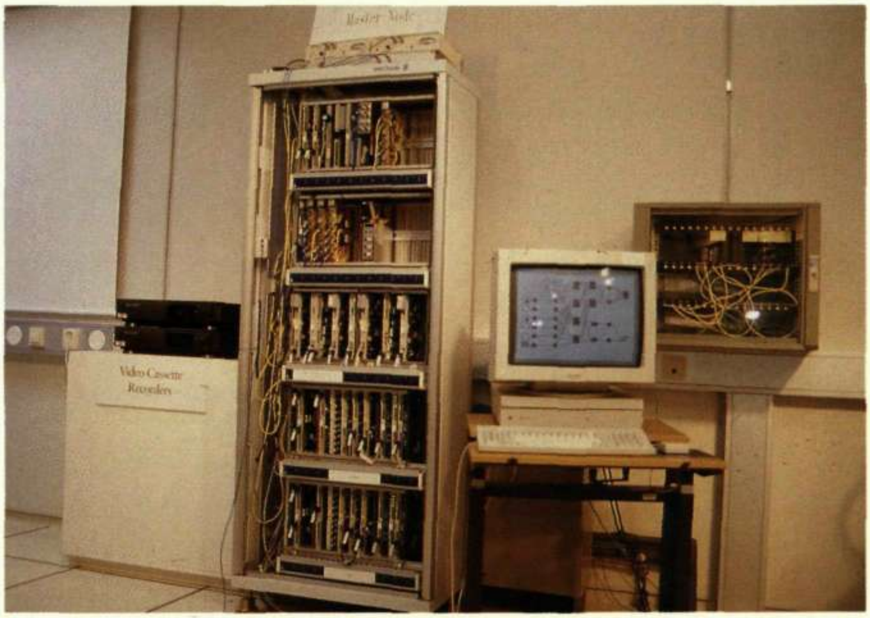


Fig. 4A
Optic switch of the Oscar demonstrator also showing the control workstation, to the right, and two service providers in the form of video recorders to the left

includes many demonstrator projects in which new hardware, often from different manufacturers, is set up to show interwork and to display interesting functionality. These systems are invariably computer-controlled, usually by workstations, and it is important that software development does not delay the project or becomes too costly.

One such project is RACE Oscar (R1033).[9] The aim of this project is to identify how and when optical switching should be used in future broadband networks and to develop the corresponding optical switching technologies. A demonstrator has been built by Ellemtel based on optical components from Ericsson Telecom. This demonstrator consists of:
- A fully transparent optical switch
- Two service providers in the form of video cassette recorders
- Three subscribers, each having two video monitors and a video camera.

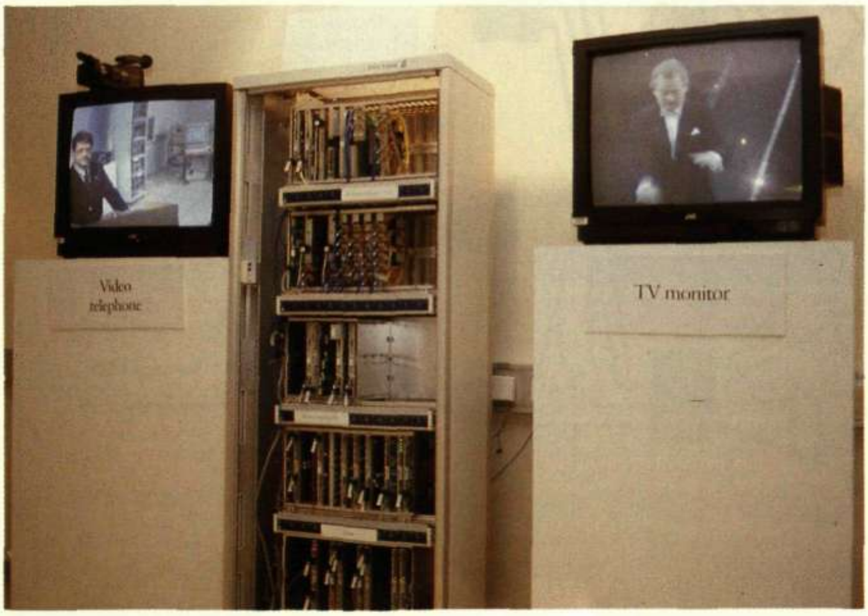


Fig. 4B
One of the subscribers or nodes of the Oscar demonstrator. The node contains two TV monitors and a video camera. In this figure the monitor on the right has been through-connected to one of the service providers, and the monitor to the left is used together with the camera as a video telephone

A photograph of the demonstrator is shown in Fig. 4. The demonstrator was first shown at the European Conference on Optical Communication, ECOC, in Paris, in 1991. It was necessary to develop a user-friendly control system that gave a good overview of the switch paths that had been set up. Fig. 5 shows an example of the workstation screen. Selection of paths is done entirely by means of a mouse and menus.

This control system was written by one person in Erlang in about two months. The work included programming the hardware interface, and the code volume was about 3000 lines. It was natural that switch points should be described as processes and switch connections as messages.

Another demonstrator project is RACE Biped (R1056) in which no less than 11 different partners participate. The aim of Biped is to form a basic business IBC (Integrated Broadband Communications) demonstrator in an evolutionary hybrid network environment by interconnecting subsystems from three other RACE projects: Business CPN (R1011), Access (R1030) and Atmospheric (R1014).

These projects cover private networks, access and switching functions. The Biped demonstrator is thus a "complete" end-to-end IBC system including the standard reference points Sb and Tb. The demonstrator – shown in Fig. 6 – has been set up at Ericsson in Madrid.

Control of this heterogeneous system is a complex matter. A proposal that was

rejected was to control the different pieces of equipment by a number of PCs connected over an LAN. Instead it was decided that a workstation should be used and that the program should be developed in Erlang. Erlang was chosen after an evaluation, because
- it was by then (end of 1991) a well tested technology
- it was easy to create graphics
- it was easy to connect Erlang to hardware and to other software.

The original plans were that a first restricted version of the software should be operational by mid-92. In fact, it became available as early as April. The software was developed in an incremental fashion, i.e. after the main system had been created, new interfaces were developed as new pieces of equipment arrived.

The entire system now consists of about 7000 lines of Erlang and about 2000 lines of C, developed in part at Ellemtel and in part at Ericsson, Madrid. This development work has required about 1.5 man-years.[15]

Conclusions

Working with Erlang encourages a "hands-on" attitude to program development. It allows professional software engineers to cover the entire spectrum from high-level systems design to concrete pro-

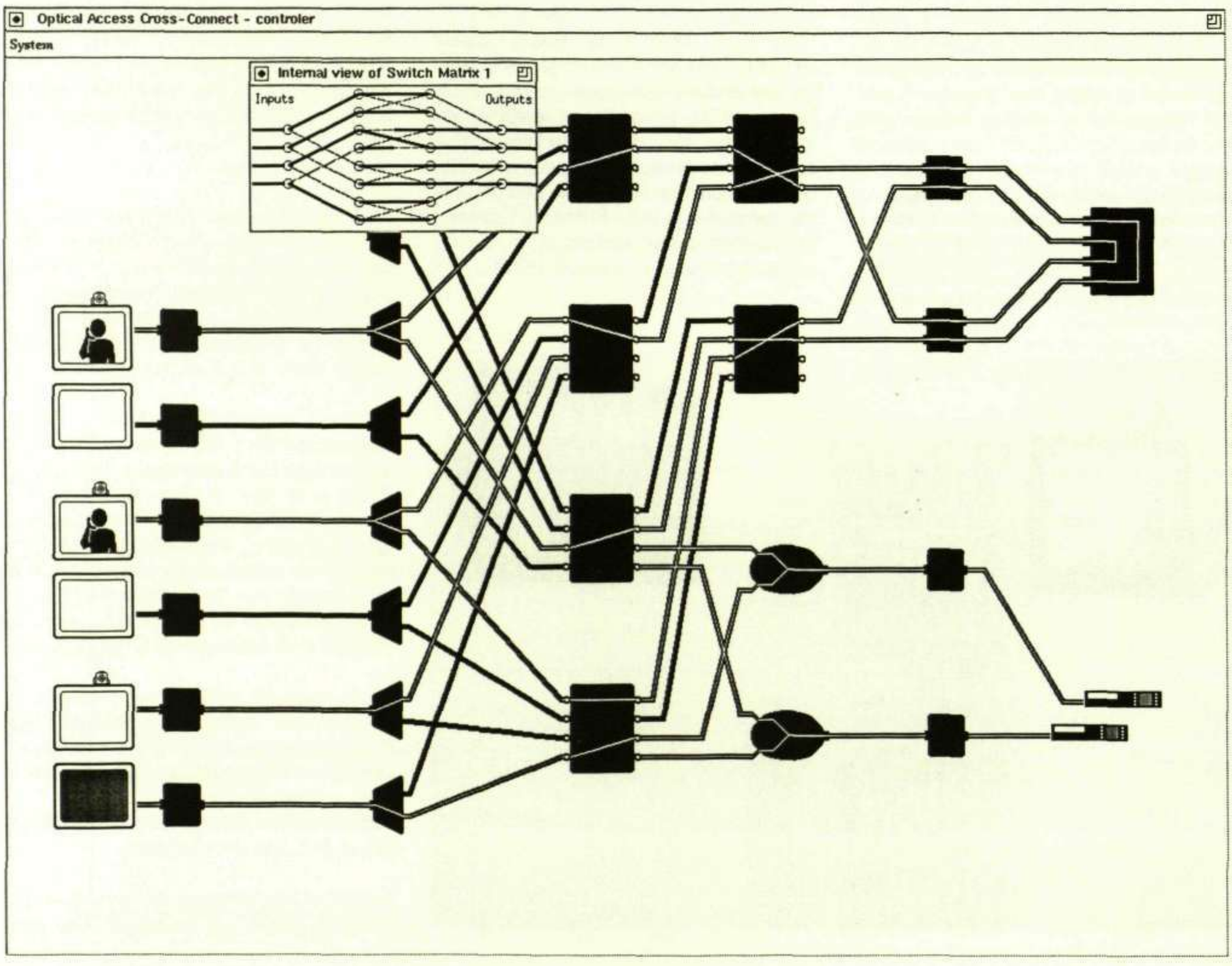


Fig. 5
A screen dump of the workstation when running the Oscar demonstrator. The picture gives an overview of the optical network with connected paths highlighted. The subwindow at the top shows the internal view of one of the switch matrices

Fig. 6
The Biped demonstrator build up at Ericsson S.A. in Madrid, Spain – the first end-to-end IBC system demonstrator within the RACE programme

gram design close to the telecommunications hardware. High productivity means small organisations.

Present Erlang implementations are fast enough for PABX applications, and faster implementations are already being tested in the laboratory. The goal is that systems developed in Erlang should not be slower or require more expensive computer equipment than similar systems developed by means of conventional languages like C++.

Erlang being a new language is obviously not itself a standard. However, every effort has been made to ensure that Erlang implementations can cooperate with standard software. Erlang implementations have thus been ported to commercially available operating systems like UNIX, MS-DOS, Macintosh and to real-time operating systems like VxWorks and QNX. Mechanisms have also been added to Erlang to make it easy to communicate with other software. A good example of this is the Interviews-based interface to the X-window system which has been written in C++.

Erlang is a programming language – an implementation technique. It does not impose an architecture but it gives system designers the elements to create architectures suited to specific problem domains. These can be as diverse as photonic switching, control of switches for cordless telephony or ASN.1 compilation.

The Erlang programming language and its implementations have now reached such a level of maturity that Ericsson has decided to productify the language as a core technology for future system products. A recently formed company will market this product and related tools.

References

1. Ahlberg, I., Bauner, J.-O. and Danne, A.:*Prototyping Cordless by Using Declar-ative Programming.* XIV International Switching Symposium. Yokohama, 1992.
2. Armstrong, J. L., Elshiewy, N. A. and Virding, S. R.: *The Phoning Philosophers' Problem or Logic Programming for Telecommunications Applications.* Third IEEE Symposium on Logic Programming, Salt Lake City, 1986.
3. Armstrong, J. L. and Williams, M. C.: *Using Prolog for Rapid Prototyping of Telecommunications Systems.* Seventh International Conference on Software Engineering for Telecommunication Switching Systems, Bournemouth, 1989.
4. Armstrong, J. L. and Virding, S. R.: *Erlang – An Experimental Telephony Programming Language.* XIII International Switching Symposium, Stockholm, 1990.
5. Armstrong, J. L., Virding, S. R. and Williams, M. C.: *Erlang User's Guide and Reference Manual, Version 3.2.* Ellemtel Utvecklings AB, 1991.
6. Armstrong, J. L., Däcker, B., Virding, S. R. and Williams, M. C.: *Implementing a Functional Language for Highly Parallel Real-Time Applications.* Eighth International Conference on Software Engineering for Telecommunications Systems and Services. Florence, 1992.
7. Armstrong, J. L., Virding, S. R. and Williams, M. C.: *Use of Prolog for Developing a New Programming Language.* International Conference on the Practical Application of Prolog. London, 1992.
8. Armstrong, J. L., Virding, S. R. and Williams, M. C.: *Concurrent Program-ming in Erlang.* Prentice Hall. 1993.
9. Buhgard, M., Granestrand, P., Johansson, S., Lagerström, B., Lindblom, M. and Thylén, L.: *Photonic Switching in High Capacity Transport Networks.* XIV International Switching Symposium. Yokohama, 1992.
10. Bâge, G. *The Programming Language EriPascal.* Ericsson Information Systems AB, 1984.
11. Däcker B., Elshiewy N., Hedeland P., Welin C. W. and Williams M. C.: *Experiments with Programming Languages and Techniques for Telecommunication Applications.* Sixth International Conference on Software Engineering for Telecommunication Switching Systems, Eindhoven, 1986.
12. Eriksson, D., Persson, M. and Ödling, K.: *A Switching Software Architecture Prototype Using Real-Time Declarative Language.* XIV International Switching Symposium. Yokohama, 1992.
13. Foster, I. and Taylor, S.: *Strand - New Concepts in Parallel Programming.* Prentice Hall, 1990.
14. Hemdal, G.: *AXE 10 – Software Structure and Features.* Ericsson Review *53* (1976):2, pp 90-99.
15. Monfort, F.: *Control Switching Implementation of the BIPED Demonstrator.* Second Australian Conference on Telecommunications Software. Sydney, 1993.
16. Steedman, D.: *Abstract Syntax Notation One (ASN.1) The Tutorial and Reference.* Technology Appraisals. 1990.
17. Wickström, C.: *Processing ASN.1 Specifications in a Declarative Language.* Eighth International Conference on Software Engineering for Telecommunications Systems and Services. Florence, 1992.

New Technology for Prototyping New Services

Kerstin Ödling

KERSTIN ÖDLING
Ericsson Business Networks AB

Software development is time-consuming, and lead times from a customer's request to the finished product is often long. Frequently, the resulting product also provides services that the customer simply does not need or services that fail to meet his expectations. Another factor of importance to the efficiency and quality of software design is the software architecture and how it matches the function specification. The programming language, methods support and support systems are other important factors.

In order to improve efficiency and quality in software design, Ericsson has tested a way of developing new PBX applications by using software prototypes and new technology. A robust test bed for new applications has been created. It is based on well-tried processor and commercial operating system products and on stable Ericsson PBX hardware. On this test bed, a software prototype comprising about 10 % of the total functionality of the MD 110 PBX system has been programmed in Erlang.

The author describes the implementation of the prototype and the experience gained. It seems likely that the productivity in software design can be increased by more than 10 times compared with the current mainstream software technology. The prototyping tools are now being used for trying out new services.

Ericsson has made great efforts to evaluate state-of-the-art software architecture and to develop a new programming language adapted to telecom switching systems with the objective of creating a new prototyping technology. A prototype has been developed and the results of this work have been evaluated on the basis of measurements, demonstrations and tests of user products. The outcome is a useful prototyping platform for the development of new applications.

Fig. 1
Simulators are important aids for the application designer. The picture shows a basic call between two digital telephones

Defining problems

Problems in the development of current systems have been pinpointed, and work phases that can be improved have been identified. The main problem is related to the gap between specifications and code. This gap consists of two elements:

- the structural difference, which can be minimised if services stated in the function specification are mapped onto feature modules in the software architecture. This introduces a customer-oriented approach into the development process, making the design objects identical with the service packages.
- the abstraction level of the code, which can be raised by application-oriented language tools, close to natural language.

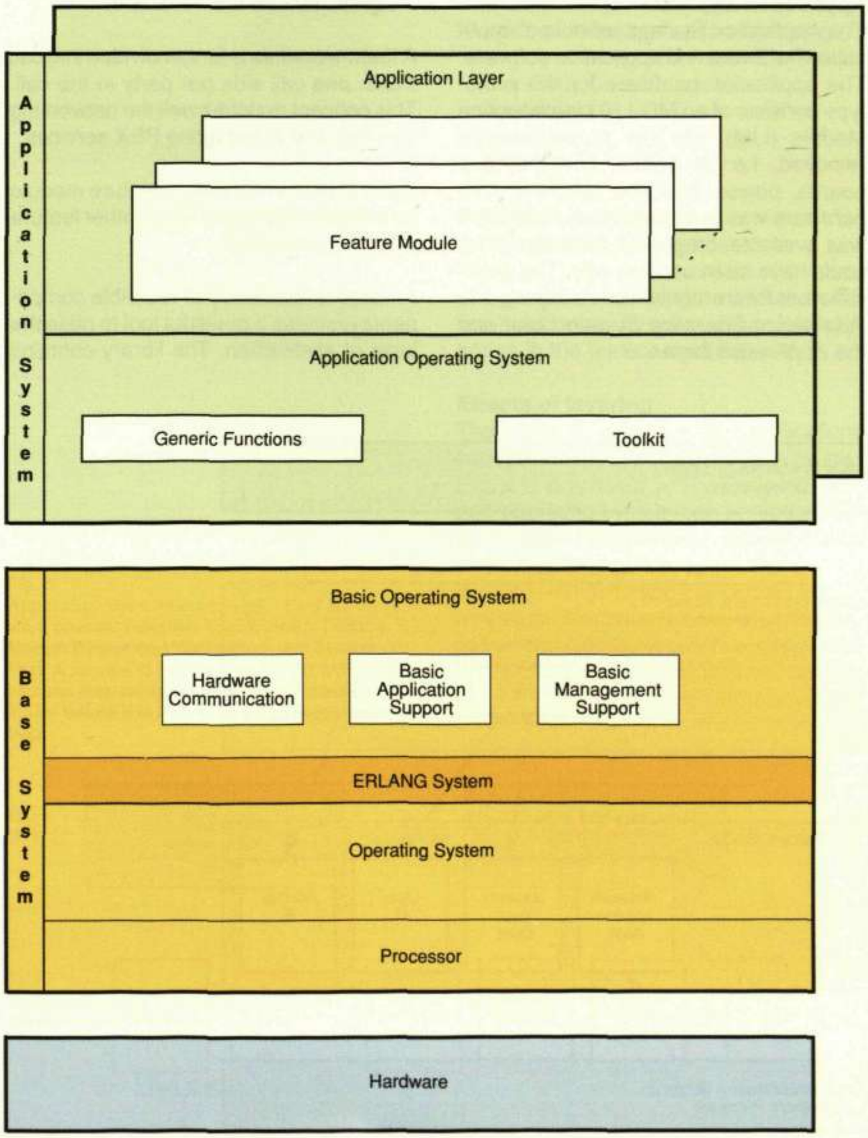
Prototyping

Prototyping is a method of solving complex problems and verifying solutions through practical experiments. Prototyping is done in cycles each of which having its own defined objectives: for example, a limited number of services which are to be fully implemented.

Software technology base for making prototypes

The software technology for prototyping, established as a part of the initial work, has proved to be a stable base for the

Fig. 2
The layered software consists of two major parts: the Base System and the Application System. The Base System in turn consists of the Basic Operating System, providing basic support to the Application System, the Underlying Operating System and the Processor System. The Application System consists of the Application Operating System Layer and the Application Layer. The Application Operating System contains a Tool Kit and Generic Functions. The Application Layer is structured in Feature Models, mapping the function specification structure of the services



creation of new prototypes. These prototypes will represent real systems, although limited in functionality. They will have the real-time characteristics and the software quality required for large-volume products.

Specification structure and methodology

The initial task when designing a prototype is to write the function specification. Specification structures and methods have been studied. A large number of specifications have been available from different standard organisations (ETSI, ECMA) in various forms similar to the so-called 3-stage method [1], [2], [3] specified by CCITT. These specifications are computer-stored and will soon be distributed on suitable computer-readable media. Since rewriting them requires a great deal of unnecessary work, it has been decided that specifications received in a form similar of the so-called 3-stage method are to be used in their original form. For entirely new applications, the specifications will be written according to a simplified 3-stage method. A template – with automated links from the overview to the document and from the table of contents to each paragraph in the document – supports the writer of the specification.

System architecture

The system architecture is layered, Fig.2, and divided into two major parts, one called *Base System* and the other *Application System*.

Base System
The Base System provides a common system base on which several types of applications can be generated. Typically, one application is Telephony and another Management.

The Base System contains the *Basic Operating System*, the *Underlying Operating System* and the *Processor System*.

Basic Operating System
The Basic Operating System provides an execution environment for applications written in Erlang and for primitives and general functions required by the applications systems, e.g.

- System start and restart
- Database storage and retrieval
- Generic functions for device allocation/deallocation
- *Generic functions for hardware drivers*
- Error recovery
- Hiding of software distribution on CPUs and system configuration

The Basic Operating System raises the abstraction level of programming and gives an essential contribution to the fault-tolerant system aimed at. It also stores information about the resources allocated to the application processes, and about which processes are linked together in each transaction. If a fault occurs, either due to a programming error or a hardware malfunction, the Basic Operating System "kills" the processes involved in the transaction and frees the resources allocated. This results in a very robust test environment in which hardware faults and software errors only affect the transactions in which they occur.

Underlying Operating System and Processor System
The Underlying Operating System and the Processor System are commercial products such as UNIX on Sun SPARCS or VxWorks on Force CPU 40.

Application System

The Application System consists of application hardware and application software. The application hardware for the prototype consists of an MD 110 Line Interface Module (LIM) with the processor part removed, i.e. of switch, line interface boards, power units and cabinets. This hardware was selected merely because it was available; other PBX hardware units could have been used as well. The application software contains two layers, the *Application Operating System Layer* and the *Application Layer*.

Application Layer
The application layer contains design objects in the form of *Feature Modules* and *Application Library Modules* which, in their structure, must closely map the application function specification structure and service packages.

The feature modules are further subdivided into:
- User modules, handling access-independent parts of the service, i.e the traffic control part.
- Access modules, handling terminal characteristics and origination/termination of call sessions.
- Driver modules, encoding and decoding signals to/from the hardware.

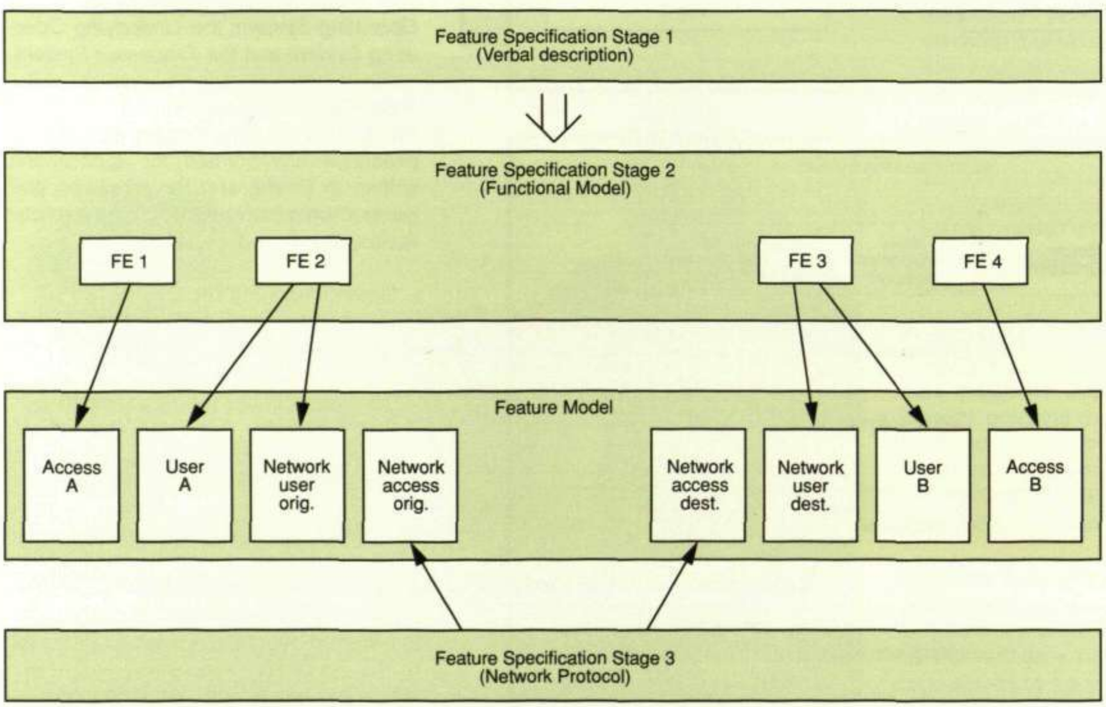
A feature module is further divided into call sides: one call side per party in the call. This concept matches well the networking services and stand-alone PBX services.

Basic Call is a mandatory feature module for telephony on top of which other feature modules are added.

An application library of reusable components provides a powerful tool to raise the level of abstraction. The library contains

Fig.3
The Call Model, e.g. the run-time structure, and the internal structure of the feature model are identical. The mapping from specification to code is shown by arrows. For a service internal to a node, there is a set of access and user modules for each side of the call. If the service is to work over the network, additional Network Access Modules are needed to carry the network protocol. If the service has additional call cases due to the network, these are implemented in Network User Modules Definitions:

A Originating side of a call
B Terminating side of a call
FE Functional Entity
Orig Originating
Dest Destination



functions which are frequently used when services are designed. These functions are identified in the specification phase, which is run through for each new application.

Application Operating System
The Application Operating System provides support functions for the application layer in order to avoid duplicating code in several different services and to raise the level of abstraction in application programming, by hiding as many implementation details as possible from the application designer. It also contains *Generic Functions* and a *Tool Kit.*

The Tool Kit provides general-purpose functions for the application layer, e.g.

- Inter Party Communication
- Connection
- Queue
- Timing
- Call History
- Number Analysis
- Configuration Management

The generic functions provide the execution mechanism for user and access programs in the feature modules.

Effects of layering
The effect of layering is that applications become independent of the operating system and processor used. The processor and operating system can be updated in pace with technological developments without affecting the applications. The layering within the application part means a raise of the conceptual level of the language and hides computer concepts from the application designer who will be able to concentrate on "what to .." instead of "how to .." aspects.

Run-time structure

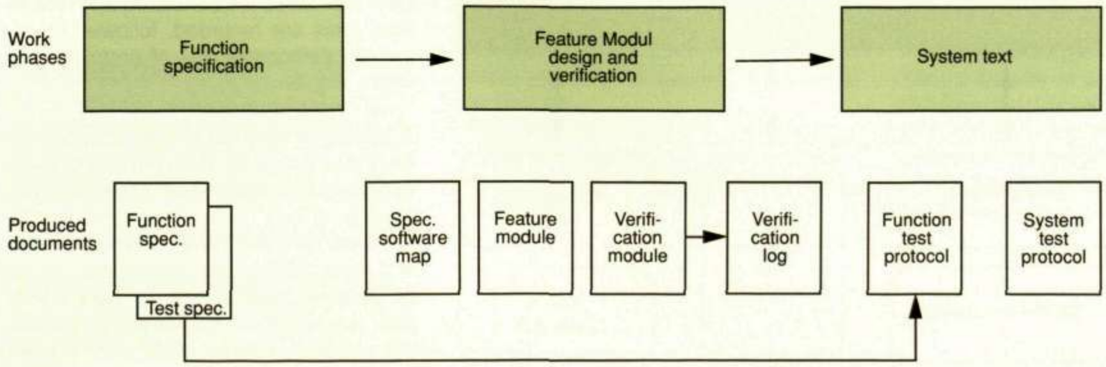
A standard-process concept has been chosen to model the run-time structure of the application, i.e. the *Call Model.* Our ambition has been to structure the Call Model in the same way as the Functional Model of the specification is structured. There is also a need to model line-terminal devices. This has led to the assignment of one process for:

- Each hardware device (driver process)
- Each type of line (access process)
- Each party in a call (user process)

The different sides of the call are represented as different sets of entities already in the functional model created in the specification work. It has therefore been found suitable to choose a split view of the call model, dividing a call into call sides with their own sets of processes, one for each party in the call. When adding a third party - to establish a conference call - a third call side with its own set of processes is created. Also, a common user process, called service user, is then created reflecting the multi-party connection. An apparent advantage of using call sides is that the total number of states is significantly reduced.

In many cases, split call control requires negotiation/communication between the call sides before decisions are made. Communication between the sides is supported by a high-level protocol which

Fig. 4
Application work methodology. There are three work phases: Function Specification, Feature Module Design and Verification, and System Test. A service to be prototyped is in turn specified and then designed and verified in a feature model before it is turned over to the system test team



hides messages passing between processes from the programmer.

Application work methodology

The run-time structure and the feature-module structure are identical. These structures closely match the specification parts, stage two and stage three, Fig. 3. The close matching of code and the run time structure to the specification structure significantly reduces the number of steps in the design process and the need of documentation, and – consequently – the amount of work to be done. The methodology and the corresponding documents produced during the different work phases are shown in Fig. 4. The actual methodology for making the prototype is reduced to three work phases: Function Specification, Feature Module Design and Verification, and System Test. Having done the work in the Feature Module Design and Verification phases is tantamount to having performed a conventional function test of the service, except for the fact that the hardware is simulated.

Very few persons need to be involved. One person does the Functional Specification phase as well as the Design and Verification phase. This is made possible because the services are contained in feature modules. System Test, which is done by a separate test team using the appropriate hardware, includes service interaction and capacity tests.

Programming language

It is well known that the use of declarative languages, such as Prolog, for programming conventional sequential applications results in small programs and significantly increased programmer productivity. A series of experiments have been performed within Ericsson to determine whether such languages can also be used to implement highly parallel real time applications and, if so, whether similar productivity gains can be made. A result of these experiments is Ericsson's development of the declarative programming language Erlang[1], with concurrency and error recovery built into it.

Implemented services

Several services were chosen as test objects for evaluating the prototyping technology. These services – implemented according to the function specifications for Ericsson's MD110 PBX – are:

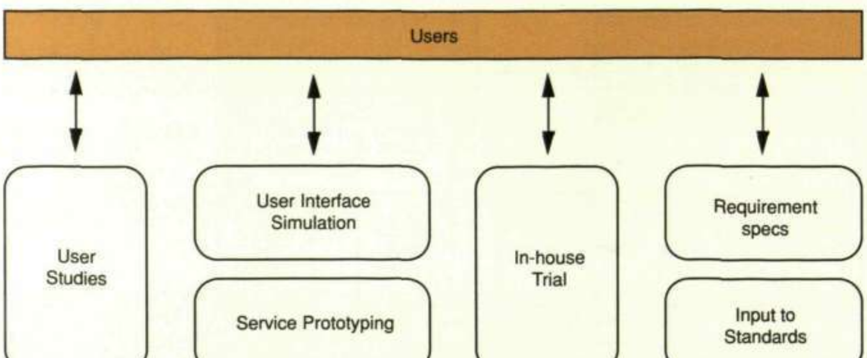
- Basic Call
- Basic Network Call
- Basic Cordless Call
- Calling Line Identification
- Three Party Services
- Call Forwarding
- Operator Extending
- Call Completion on Busy/No Reply
- Operator Recall
- Intrusion

Overall methodology

The new technology has proved to be very efficient and has enabled us to make real-life prototype tests of new services involving users and to revise specifications on account of the users' opinions. Service prototyping has been made feasible. The methodology involves the user from the start; even in studies of new applications. User interfaces are simulated and results from tests are recorded, followed by in-service demonstrations of prototypes to users, Fig. 5.

A prototype with some new applications has been designed, installed and put into service at Ericsson's Bollmora and Sundbyberg offices. User reactions to the installed services have been evaluated and the specifications revised accordingly. Updated applications will be put into service and user reactions to the updates will also be evaluated. Prototyping is used

Fig. 5
The needs and behaviour of service users (*in the upper part of the picture*) are studied.
New service ideas arise which are turned over to a team that makes simulations of the user interface, and to another team that makes a prototype for real-life tests.
These activities result in requirements specifications for services which meet customer needs and have user-friendly interfaces



as a tool for creating specifications for new types of applications.

Experience

Application designers
The layered approach to software architecture combined with the close mapping of the software architecture to the functional structure of the specification makes it possible for one person to perform the design and implementation of a service. He/she may be in charge of both the "function specification" and the "feature module design and verification" work phases. This will allow application designers to focus on user requirements and user behaviour. Lower layers in the architecture are the responsibility of system designers.

Increase in design efficiency
The close mapping of the software architecture to the functional structure makes the work model very simple. As a consequence, the number of documents to be produced is significantly reduced. Some of the documents in the prototype are generated automatically. Measurements of design efficiency compared with present systems indicate an increase in efficiency in the order of 10.

Easy and confident planning
The number of persons involved in function specification and in design and verification of a specific service is reduced to one, which offers obvious advantages. Annoying waiting times are eliminated, lead times are reduced and planning becomes more accurate.

Stimulating design and verification
The main factors that make it stimulating and easy to design a feature module are:

- The written text of the function specification corresponds to the code in the corresponding feature module, which improves the understanding of the code or service
- Programs can be made small and surveyable with language features such as *matching, list handling* and *recursive functions*
- The design is incremental and interactive and allows structured growth
- Patching is not needed since programs can be recompiled on the fly
- Repeated verification of parts of a complete service is "automatic", i.e. merely by activating the test file
- Data can be displayed at a high symbolic level
- Fewer documents have to be written

Performance versus CPU requirements
Erlang was first implemented as an interpreter written in Prolog. This was a useful way to develop and experiment with Erlang, but real-time performance was unsatisfactory. The present implementation of Erlang is based on a virtual machine for which an emulator has been written in C. The compiler which generates code for this machine is itself written in Erlang. The performance of this implementation is very promising and constantly improving.

Conclusions

The architectural work has been based on a sound knowledge of applications. This made it possible to define a limited number of well-defined entities in a layered structure.

Combining this architecture – which is easy to understand and handle – with a real-time declarative language, has greatly reduced the amount of work required to design new services in telecommunication systems.

It has been possible to perform real-life tests of new services by using advanced prototypes prior to full-scale introduction of the services on the market.

These technologies, when sufficiently mature for use in products, will make it possible to move the volume of design efforts from implementation work to evaluation of the capabilities of new services and the adaptation to customer needs and behaviour.

References

1 CCITT I.130. *Method for characterisation of telecommunication services*
2 CCITT Q.65. *Detailed description of Stage 2*
3 CCITT I.310. *ISDN Network Functional Principles*
4 Ahlberg, I,Bauner, J-O. and Danne, A.: *Prototyping cordless by using Declarative Programming.* Ericsson Review *70* (1993):2.
5 Däcker, B.: *Erlang - A New Programming Language.* Ericsson Review *70* (1993):2

Prototyping Cordless Using Declarative Programming

Ingemar Ahlberg, John-Olof Bauner and Anders Danne

A declarative programming language for real-time applications has been developed by Ericsson, as part of applied research into future programming techniques. In order to test the possibilities offered by this language, a joint project has been carried out by Ericsson Radio Systems and Ellemtel.
The new programming language, Erlang, was used for developing a prototype of the control software for Ericsson's DCT900 digital cordless telephone system. The prototype system was developed by three persons in six months; their work included several rewrites of the system to find the best system structure. The system was up and running two weeks after the actual radio exchange hardware arrived.
The authors describe the implementation of the control software for the prototype system.

A declarative language like Erlang, which yields a very succinct code, gives in itself shorter lead times for system specification, design and testing. To further reduce the prototyping time, two valuable tools were developed within the framework of the prototype project: a sequence chart editor for signal flow specification and a graphical simulator for test of the control software before the cordless system hardware became available. The control software, the sequence chart editor, and the graphical simulator were all developed in Erlang.

The Erlang programming language

Erlang is a programming language that combines ideas from concurrent programming – as in Modula or Ada, for example – with ideas from declarative or symbolic programming, typified by Lisp or Prolog. Many examples show that programs written in languages like Prolog are many times shorter than their equivalents in conventional languages, such as C or Pascal. The basic idea behind Erlang is to bring this technology into the realm of industrial real-time applications.

Erlang is similar to Prolog with respect to pattern matching of arguments, and in that functions are separated into distinct clauses. Statements with clauses are similar to those in functional languages, but there are no loop constructs or goto statements – only recursions. Data structures of the language are atoms for storing fixed atomic data, numbers for storing integers and floating point numbers, tuples for grouping together a fixed number of items and lists for grouping a variable number of items. Erlang provides pattern matching in function definitions which makes it possible to write programs in a markedly declarative style – almost rule-based – as shown in the following example:

```
reverse(L) -> reverse(L,[]).
reverse([],L) -> L;
reverse([Head|Tail],L)    -> reverse(Tail,[Head|L]).
```

The function *reverse*(L) reverses a list. When called with a non-empty list, the third equation will match, and this equation will then match until the list is empty. When called with an empty list the second equation will match and the reversed list will be returned.

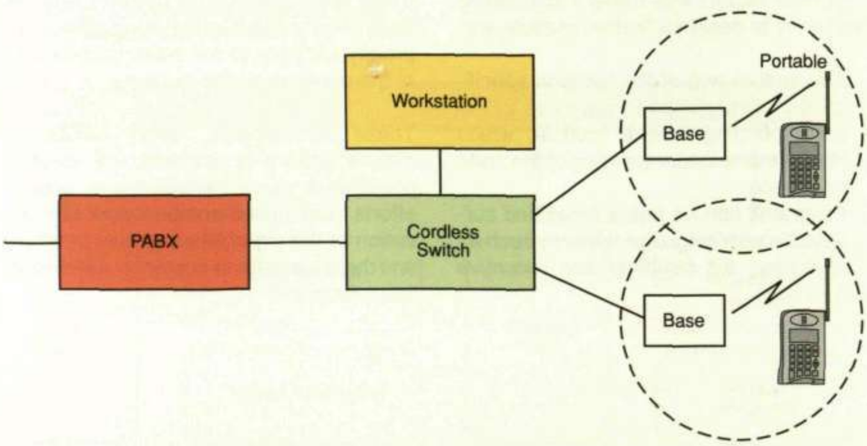
Variables are write-once-only and are used as in mathematics or in other functional languages.

Erlang handles concurrency explicitly through light-weight processes. Process sizes increase and decrease dynamically. This makes it possible to use many processes and describe the application quite regardless of implementation considerations.

Processes communicate by sending and receiving messages in a way similar to SDL. Erlang uses active signal reception where each receive statement specifies which signal it expects to receive.

Erlang uses recursion to describe iteration through the technique of last-call optimisation. This also makes for a pure style

Fig. 1
System environment with a workstation and system hardware



ANDERS DANNE
JOHN-OLOF BAUNER
Ericsson Radio Systems AB
INGEMAR AHLBERG
Ellemtel Utvecklings AB

of programming, as exemplified in the following function representing a process which holds a counter in a system:

```
counter(N) ->
   receive
      {increment} -> counter(N + 1);
      {decrement} -> counter (N - 1) ;
      {present_value,Other_process} ->
         Other_process ! {value,N},
         counter (N)
   end.
```

External communication (with hardware, for instance) is accomplished through ports which function analogously with processes. Erlang handles time through time-outs.

The *error-handling mechanisms* constitute another feature of the language. Reliability is of prime importance in real-time systems. Erlang makes it possible to trap errors and terminate and restart parts of a system without affecting the rest of it. This facility is also useful when developing and testing new functionality.

In spite of its high functionality, Erlang is a very small language and the normal course given to teach it takes about five days. As a final task, the course members develop a small but working POTS. This course has also been given at various universities as an example of declarative concurrent programming.

The DCT900 radio exchange

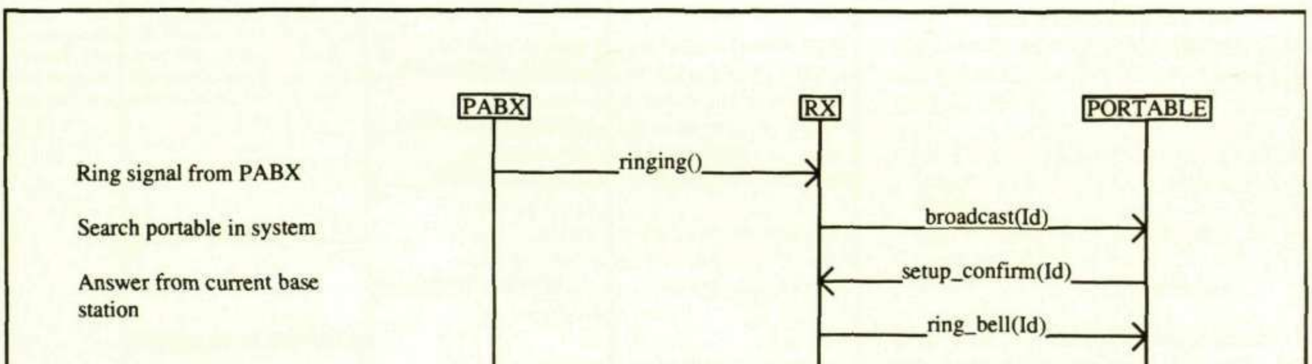
The DCT900[4] system is similar to a DECT[5] system based on Multicarrier/ Time Division Multiple Access/Time Division Duplex (MC/TDMA/TDD) technique. It is designed for a high density of users with handsets offering voice quality comparable to that of wired extensions. A cordless radio subsystem consists of a radio exchange, base stations and portable phones.

The *radio exchange cabinet* used in our experiment was equipped with four printed board assemblies: the CPU, the Line Termination Unit (LTU) for connecting PABX lines (telephone extension cables), the Speech Processing Unit (SPU) for call set-up, and the Cell Controller Unit (CCU) for connecting and handling Cordless Portable Parts (CPP) via two base stations as tranceivers.

System environment

The Erlang program development environment runs on a Sun/UNIX workstation. In this prototype system the same workstation also controls the cordless exchange. Thus, the functions of the host and the target computer are combined, which provides a very efficient development environment, Fig. 1.

Fig. 2
Specification of a call from a PABX to the portable at the highest system level



The DCT900 radio exchange is connected to the workstation via an RS232 serial line. The DCT900 CPU software contains a transparent mode option. When this option is set, calls originating from the CCU (portable phone) or the LTU (PABX) are not handled by the CPU software but are sent out on the RS232 serial port of the CPU. A call is thus treated by the Erlang software system, and signals are sent back over the serial line and the CPU to the appropriate unit.

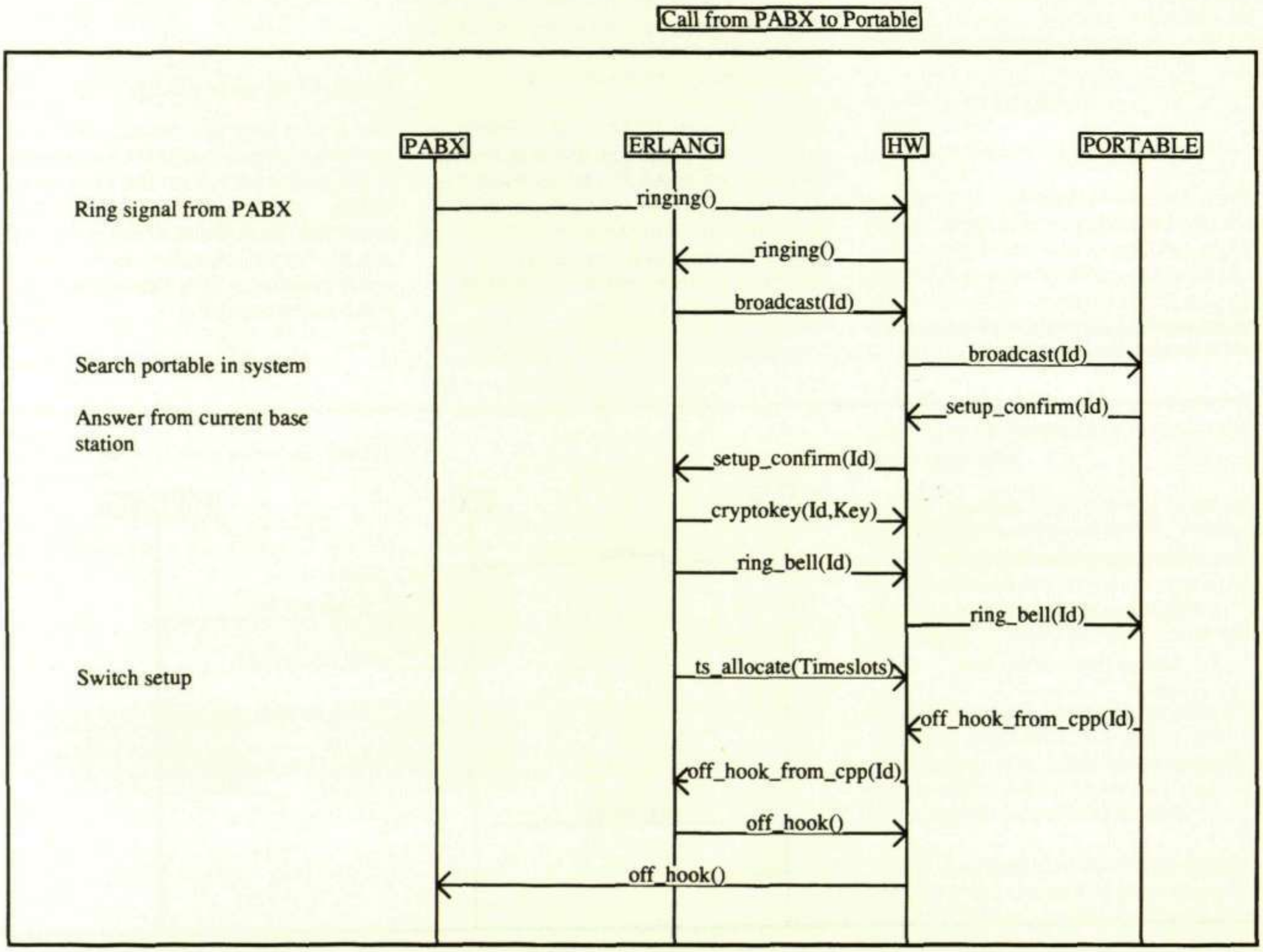
Program specification and implementation

The system was structured on the basis of documentation that describes the internal and external signalling in the DCT900 system. A software tool was developed and used for handling sequence flow charts.

Sequence Chart Editor
Sequence charts can be used to specify and illustrate the behaviour of real-time systems. A sequence chart editor, *sc*, was developed in Erlang using an X-Windows library called *pxw*. The *sc* has the following main features:

- Processes are represented as vertical lines
- Sequences of messages between processes can be drawn with the mouse. Messages are represented as horizontal arrows
- Process states can be inserted
- Functions represented as sequences are saved in a database

Fig. 3
The radio exchange split into control software and hardware



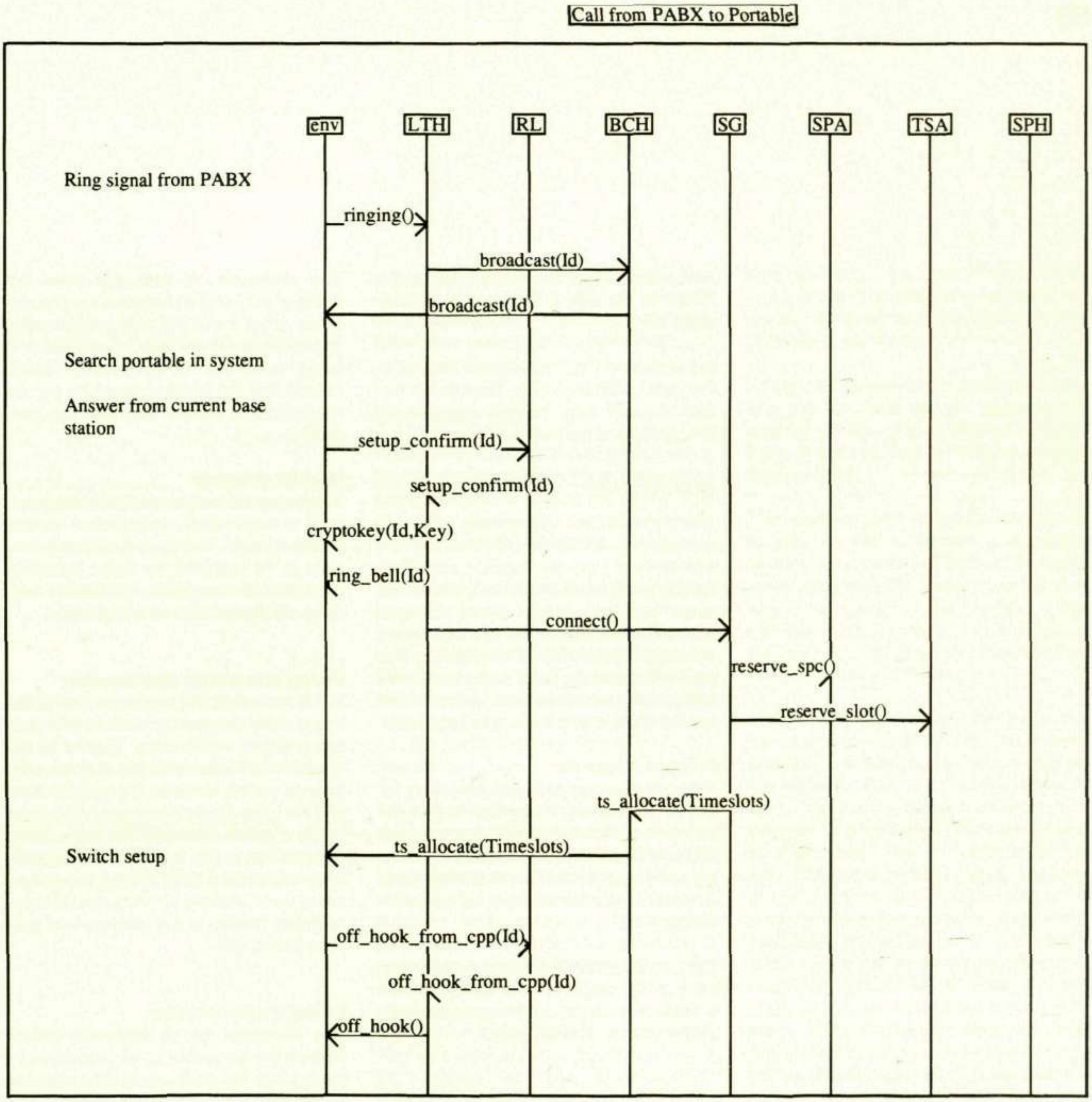


Fig. 4
Internal process interaction in the Erlang control software. The rest of the system is hidden in the "env" process. The sequence chart shows the same call to a portable as that illustrated in Fig. 2

- The process structure can be hierarchical. A process can be split into a set of subprocesses. A set of processes can also be combined to form a compound process. Thus, it is possible to apply both a top-down and a bottom-up method of working
- Different views of the process hierarchy can be specified. It is possible to work in out-line mode, showing sequences at various levels of detail

The feasibility of generating Erlang code automatically from the sequence chart database is being studied.

System specification

Different function and signal specifications for the DCT900 were interpreted and represented as sequence chart diagrams using the *sc* editor. The following functions were illustrated:

- Call from a PABX to a cordless portable
- Call from a cordless portable to the PABX
- Disconnection from the portable
- Handover during a call
- System start

The functions were presented at a high level where the interface between the Erlang switching software and the radio exchange was easily recognised. No details of the internal structure and behaviour of the Erlang system were revealed in these views of the system. The system charts were reviewed by the DCT900 developers.

The chart in Fig. 2 illustrates a call from a PABX to a cordless portable at system level.

In the chart in Fig. 3 the signals internal to the radio exchange (RX) have been further divided into control software (-ERLANG) and hardware (HW).

Architecture

The architecture of the Erlang switching subsystem - called ERLANG in Fig. 2 - was derived from the function specifications in the form of sequence charts. This is an iterative process where changes have to be made when different functions are superimposed on the architecture. For example, message parameters have to be analysed in order to determine whether or not the chosen architecture is applicable.

Process structure

A map of the processes are necessary for further breakdown. Assumptions have to be made as to which processes are static (i.e. created at system start and existing permanently) and which are dynamic (created to exist temporarily, for a speech connection).

It is also a matter of whether the process will exist as only one instance, or whether several instances of the same process will be running concurrently.

The programming language does not make any distinction between the process types, since these are only programming abstractions. There are no problems in using hundreds of processes, which means that the process structure can be modelled in a natural way with few constraints.

Module structure

Processes will be grouped in Erlang modules. In most cases, one process resides in one module. The control software consists of 18 modules for traffic handling, hardware communication, database handling, configuration and initialisation.

Erlang subsystem specification

In the next step, the sequence charts are made more detailed by means of the chosen process architecture. Thanks to the hierarchical implementation of *sc* it is possible to switch between the system level and the Erlang subsystem level. The chart in Fig. 4, which illustrates the same traffic case as that in Fig. 2, shows internal process interaction. Only Erlang processes are shown. The rest of the system is automatically hidden in the environment process called *env*.

Erlang implementation

The complete set of sequence charts defines the behaviour of all processes in the system. For each process the relevant information is collected from all sequence charts. This defines, completely, how the process interacts with its environment.

Process states are then added to the sequence charts for the process in order to define the state machine. Each state where the process waits for one or more signals will be realised as a *receive* statement. When the state machine has been defined, the translation into Erlang is easy and straightforward.

The example in Fig. 5 is part of the traffic control program that handles call set-up. This part handles the start of an incoming call from the PABX side.

Project data

The traffic control modules consist of about 400 lines of Erlang code. The same functionality implemented in Modula-2 requires about 2500 lines. By using

Fig. 5
Part of the traffic control program handling a call from a PABX to a portable

```
-module(lth)                          %% line termination handler
%% wait 2000 ms for confirmation
waitformore(idle, Ptn, Ltc_address) ->
receive
        ringing -> broadcast_and_wait(Ptn,Ltc_address,2000,2)
end.

broadcast_and_wait(Ptn,Ltc_address,Ticks,0)
        waitformore(idle,Ptn,Ltc_address);
broadcast_and_wait(Ptn,Ltc_address,Ticks,N) ->
        send(bch, {broadcast,Ptn}),
        receive
                {setup_confirm, Ccc_address} ->
                        Crkey = get_db(cpp,{ptn,Ptn},
                        send(rx, {Ccc_address,{cryptokey,Ptn,Crkey}}),
                        send(rx, {Ccc_address,{ring_bell,Ptn}}),
                        send(Sgid, {connect,Ptn,Ltc_address,Ccc_aress,
                                start_speech_handling}},
                        waitformore(to_cpp,Ptn,Ltc_address)
                after Ticks ->
                        broadcast_and_wait(Ptn,Ltc_address,Ticks,N-1)
        end.
```

Erlang, the amount of source code was reduced by a factor 6. This agrees with other applications for which comparisons have been made with C++ and Ada.

The Simulator

Part of the prototype is a graphical simulator on the workstation which gives a pictorial presentation of the hardware. This simulator was used to test the software before the hardware became available.

To confirm that the implementation of the control software was correct, we implemented our signalling charts in Erlang and built a graphical simulator of the hardware that described - at a conveniently high level - the behaviour of the radio exchange. This enabled us to check that our programs behaved according to our intentions.

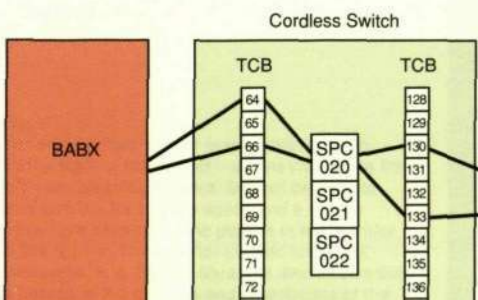
For example, the function that handles a call from a PABX to a cordless portable can be simulated by pressing the RING button of the line interface, Fig. 6. The simulator sends a signal *ringing* to the control software which responds with a signal *broadcast*. The corresponding handset is located to a specific base station. A signal *setup_confirm* is sent back to the control software, which selects a switching element for the call, allocates the necessary time slots and then sends the following signals for call establishment: *ring_bell*, which makes the text RING appear on the display of the handset, and *timeslot_allocate* signals, which are displayed by the simulator as lines drawn between the current line interface and a time-slot bus, then as lines drawn onwards to the switching element, to another time-slot bus and, finally, to the selected base station.

All signalling between the control software and the simulator can be displayed in another window. Fig. 7 depicts a signal trace of the call from a PABX to a cordless portable.

The handsets can be moved on the screen to simulate the movement of a person walking about with a handset. An algorithm calculates the selection of the current base station for the handset. If the handset is too far away from any one of the base stations, it will be recognised as not being active in the system.

When a call has been established and a handset is moved around on the screen, another base station may be selected by the handset. This is called *handover*. A signal *handover* is sent to the control software which deallocates old time slots and allocates new ones at the base station side. The corresponding signals are sent to the simulator, and the connections on the screen are changed accordingly. If the handset is moved too far from the base stations and the handset remains out of reach for ten seconds or more, a signal *unusual_clear* is sent to the control software, which deallocates time slots, etc, and clears the call by sending *release* signals to the simulator. All the connections related to the call are erased on the screen.

Fig. 6
The interactive simulator shows the functionality of the control application. A connection to one of the portables is set up through the switch



The Iconoscope

When the radio exchange was delivered, the sending and reception of signals from and to our control programs were switched over to the radio exchange itself instead of to the simulator. To effect this, a C program was developed that took care of the lower layers of the communication protocol towards the radio exchange. In order to graphically trace call set-ups etc. in the radio exchange, the simulator was changed into an iconoscope – a passive unit that only receives signals sent between the hardware and the control software in the workstation. The iconoscope takes appropriate actions that result in changes in the graphical presentation of the radio exchange.

Erlang Simulator	Address	Signal
<---	{0,3,1}	ringing
--->	{0,1,8}	{broadcast,[3,2]}
<---	{0,1,2}	{setup_confirm,[3,2]}
--->	{0,1,2}	{cryptokey,[3,2],[]}
<---	{0,1,2}	{valid_speech,[3,2]}
--->	{0,1,2}	{ring_bell,[3,2]}
--->	{0,3,8}	{ts_allocate_pair_ltc_spc,{0,3,1},64,65}
--->	{0,2,8}	{ts_allocate_both_pairs_spc,{0,2,0},129,128,65,64}
--->	{0,1,8}	{ts_allocate_pair_ccc_spc,{0,1,2},[3,2],128,129}
--->	{0,2,0}	start_speech_handling

Fig. 7
Signal trace of a call from a PABX to a cordless portable in the interface between the control software and the simulator

Later on, the iconoscope has been further developed to receive signals sent between the hardware of the radio exchange and its control programs. (These programs are written in Modula-2 and run on the computer of the radio exchange.) Switching is no longer performed in the workstation; only presentation. This has been done in order to show the activity in the cordless system without degrading performance and robustness.

The iconoscope is similar to the simulator, but it is run concurrently with the hardware to show how this is working; when handover is performed, for example. While the simulator has been used mainly for test of the control software, the iconoscope has been used as a demonstrator of the logical parts of the radio exchange.

Conclusions

The Erlang programming language has turned out to be very convenient as a modelling tool for a cordless switch. The graphical interface available with the language is also very useful for the creation of graphical and user-friendly tools like simulators. The program volume and development efforts are considerably less extensive than those of a conventional programming language used for the same task. Execution times were also fully acceptable.

References

1. Däcker, B.: *Erlang — A New Programming Language.* Ericsson Review *70.*(1993):2.
2. Ödling, K.: *New Technology for Prototyping New Services .* Ericsson Review *70.*(1993):2.
3. Armstrong, J. and Virding, R.: Erlang - *An Experimental Telephony Programming Language.* International Switching Symposium, 1990, Stockholm.
4. Klein Wolterink, G and Åkerberg, D.: *A Business Cordless PABX Telephone System On 800 MHz Based On The DECT Technology.* International Switching Symposium, 1990, Stockholm.
5. *Digital European Cordless Telecommunications, Preliminary Standard,* 1991, European Telecommunications Standards Institute, Valbonne, France.

Low-Loss Splicing Technique for Erbium-Doped Optical Fibre Amplifiers

Wenxin Zheng

WENXIN ZHENG
Ericsson Business Networks AB

The development of the erbium-doped fibre amplifier has opened the way to an essentially transparent optical network without bandwidth-limiting electronic signal repeaters. However, if erbium-doped fibre in the amplifiers were abruptly interfaced to dispersion-shifted or conventional fibre, the transmission-loss penalty would be unacceptable. A unique technique developed by Ericsson solves this problem: splices with a loss below 0.1 dB have been made possible through the implementation of a newly developed technique for fiber fusion splicer.

The author describes the technology employed to achieve consistent low-loss splices to erbium-doped fibre and how the results have been verified.

The diffusion speed of most erbium fibre designs is higher than that of other fibres. By monitoring - in real time during the fusion process - the refractive index difference between the fibres and by controlling the thermal diffusion of the fibre core dopant, a tapered transition region between the mating fibres can be created. The resultant splices, with consistently low loss, give read-out estimates of the loss values of individual splices as a by-product. Different types of erbium fibre have been spliced to different types of singlemode fibre and dispersion-shifted fibre. Low splice loss, high strength, and a good correlation between splice loss estimates and measurements have been achieved.

Erbium-doped fibre amplifiers

A significant technological evolution is taking place in optical fibre telecommunications. Until now, the normal way of compensating for transmission loss in silica optical fibre (0.2 - 0.4 dB/km) is to employ electronic signal repeaters. Typically, the repeater consists of an optical detector followed by electronic amplifiers, timing and recovery circuits, and a laser diode that launch the signal into the next span. With such repeaters, the ultimate band-

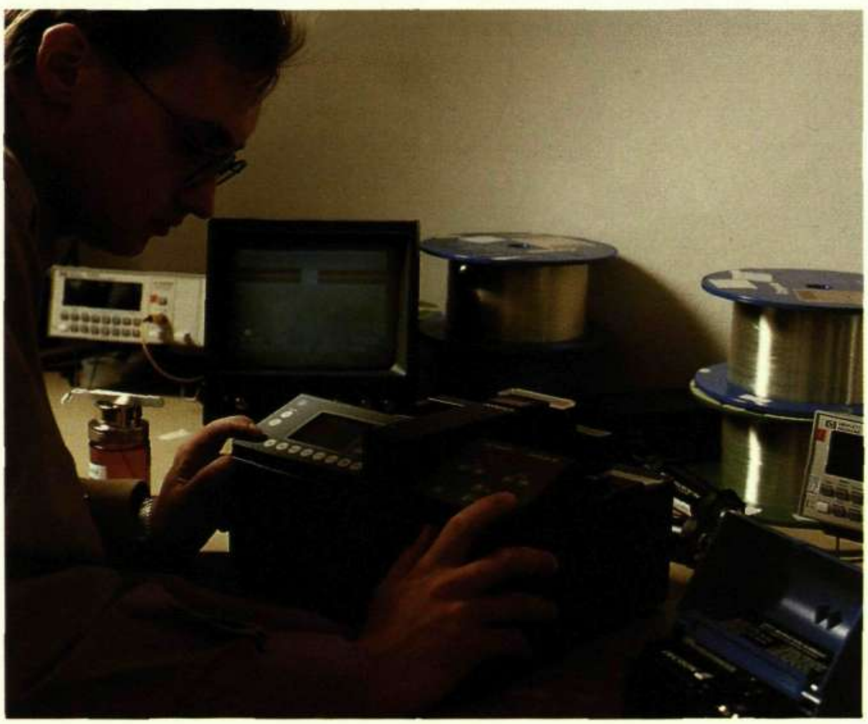
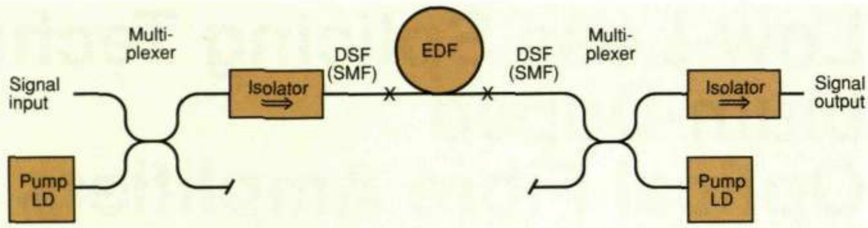


Fig. 1
The erbium fiber splicer in lab environments. To the right of the splicer – a fibre cleaver, to the left – the cleaning solvent. Behind devices to measure the loss of the splice and a screen showing a blow-up of the picture in the monitor in the splicer. The monitor is used for promt messages, e. g. type of fibre, but also shows the progress of the process and an estimate of the loss of the splice. The automatic fusion process starts when two fuse buttons are pressed simultanously

Fig. 2
Configuration of an erbium-doped fibre amplifier with bidirectional pumping

EDF	Erbium-doped fibre
DSF	Dispersion-shifted fibre
SMF	Singlemode fibre
LD	Laser diode



width of a silica fibre cannot be fully utilised because of the limitations of the electronics in the repeater.

The development of the erbium-doped fibre amplifier (EDFA) – an all-optical amplifier without the limitations of the electronics – has opened the way to an essentially transparent optical network.

The main advantages of the EDFA can be summarised as follows:

- high efficiency: 30 dB gain for about 20 mW pump power
- polarisation-insensitivity, and immunity to crosstalk
- independency of modulation format and bit-rate; ease of system upgrading.

A typical EDFA configuration is shown in Fig. 2. The EDFA consists of two diode laser pumps and two wavelength division multiplexers which couple the pump light (0.98 or 1.48 μm wavelength) and the 1.55 μm signal into the isolator and then into the erbium-doped fibre (EDF). In order to achieve higher efficiency, an EDF with a small core diameter (3 - 5 μm), but very high (about 0.25 - 0.3) numerical aperture (NA) has to be used in the EDFA.[1] However, a decrease in fibre core diameter and an increase in NA lead to a decrease in the mode field diameter (MFD) of 4 - 6 μm. This will increase the mismatch between the field distributions in the EDF and standard singlemode fibres (SMF) with an MFD of about 10 μm or dispersion-shifted fibres (DSF) with an MFD about of 7 μm, which results in great splice losses.

The intrinsic splice loss between two fibres with different MFDs can be calculated from:[2]

$$\text{loss}=20\log(w_1^2+w_2^2)/2w_1w_2 \text{ (dB)} \quad (1)$$

where w_1 and w_2 are mode-field radii for the two fibres being spliced. The calculated splice loss can be as high as 1.0 – 2.0 dB. This increase in splice loss degrades the amplifier gain, noise figure and amplifier output power.

Several successful attempts to overcome this problem in laboratories have been reported.[3-6] However, no commercially available splicers have been found to consistently solve the problem until the technique described in this article was developed.

Core material diffusion monitoring

It is well known that most singlemode optical fibres, except pure silica core fibres, consist of a pure silica cladding and a germanium-doped silica core. Besides germanium, other materials can also be doped in – or in the surroundings of – the core region, e.g., fluorine for cladding depressed fibres and erbium for fibre amplifiers. Different dopant materials and different levels of dopant concentration result in different refractive index profiles of the fibre. In general, a high concentration of the core dopant leads to high refractive index differences between the core and cladding, to a larger numerical aperture, and to a smaller mode field diameter, as explained in Box A.

With heat-treatment (by flame, filament, electrical arc, etc.), the dopant material of the core tends to diffuse into the cladding. This diffusion is related to the temperature and time of heating. The diffusion speed depends on the temperature, the dopant concentration of the core, and on the dopant material surrounding the core.[9] Given the same temperature and time of heating, different fibre designs result in different diffusion speeds in the core. For example, the higher the dopant concentration, the higher diffusion speed can be obtained. Thus, a fusion process of some duration is often used to minimise the mode field mismatch of two fibres, especially when erbium-doped fibres are to be spliced.[3-6] A typical relation between

Box A

If Δ is used to denote the relative index difference:

$$\Delta \equiv \frac{n_{core}^2 - n_{clad}^2}{2n_{core}^2} \quad (2)$$

the numerical aperture (NA) and the mode field diameter (MFD) can be expressed by [7,8], respectively:

$$\mathrm{NA} = n_{core}\sqrt{2\Delta} = \sqrt{n_{core}^2 - n_{clad}^2} \approx \sqrt{3(n_{core} - n_{clad})} \quad (3)$$

and

$$\mathrm{MFD} = \frac{2a}{\sqrt{\ln[ka(\mathrm{NA})]}} \quad (4)$$

where a is the core radius, k is the wave-number, $k=2\pi/\lambda$ is the wavelength, and n_{core} and n_{clad} are relative refractive indices of core and cladding, respectively.

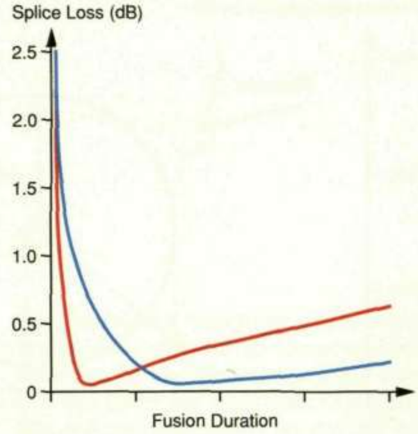


Fig. 3
Typical relation between splice loss and duration of fusion process for erbium-doped fibre when spliced to dispersion-shifted fibre

High fusion temperature
Low fusion temperature

splice loss and arc fusion time is plotted in Fig. 3, where EDF and DSF are spliced. The unit of time in Fig. 3 is determined by the fusion temperature and by the dopant material surrounding the core: it may be seconds, minutes, and even hours with electrical arc, filament, or flame splicing.[3-6] Even if only arc fusion is used, different values of the fusion current can lead to variations in the fusion time – up to a tenfold increase for reaching the minimum splice loss point.[4]

In order to obtain the lowest possible splice loss, it is necessary to measure the loss changes during the fusion process with a very stable laser source and an optical power meter. A splice loss of 0.15 – 0.50 dB can be achieved in erbium fibre splicing by manually stopping the fusion at the right time.[4-6] However, this active real-time loss measurement can only be accomplished by means of expensive laboratory equipment.

To develop a practical erbium fibre splicer with an automatic mode-field matching process, the chief thing is therefore to find a way to monitor the erbium fibre splicing process through passive splice loss measurement, without the use of a laser source or power meter.

Image processing of heated fibres

In most automatic optical fibre splicers, a video camera and an image processing system are often installed for fibre alignment. When an optical fibre is heated, the thermal radiation emitted from the fibre can be observed by the video camera and analysed by the digital image processing system. During the fusion process, the luminescence intensity from the heated fibre is different with different dopant material in silica as well as for different doping densities, as shown in Fig. 4. Thus, the Ge-doped core looks much brighter than the pure silica cladding. The light intensity profiles are also quite different for different types of fibre, as illustrated in Figs. 5, 6a and 6b.

Through the application of a band-pass filter technique in the spatial frequency domain to reduce the effect of the cladding[10], a heated-fibre index profile can be obtained from the digital image of the heated fibre. A useful relation can be obtained[10] by comparing the refractive index profiles of the cold fibre with those

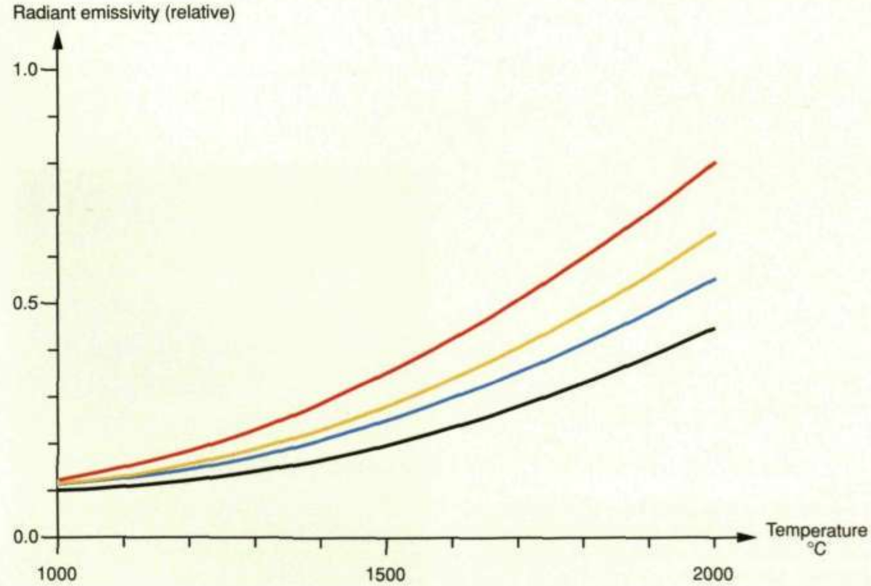


Fig.4
Thermal radiated light intensity of silica with different dopants

Er- and Ge-doped silica
Ge-doped silica
F-doped silica
Pure silica

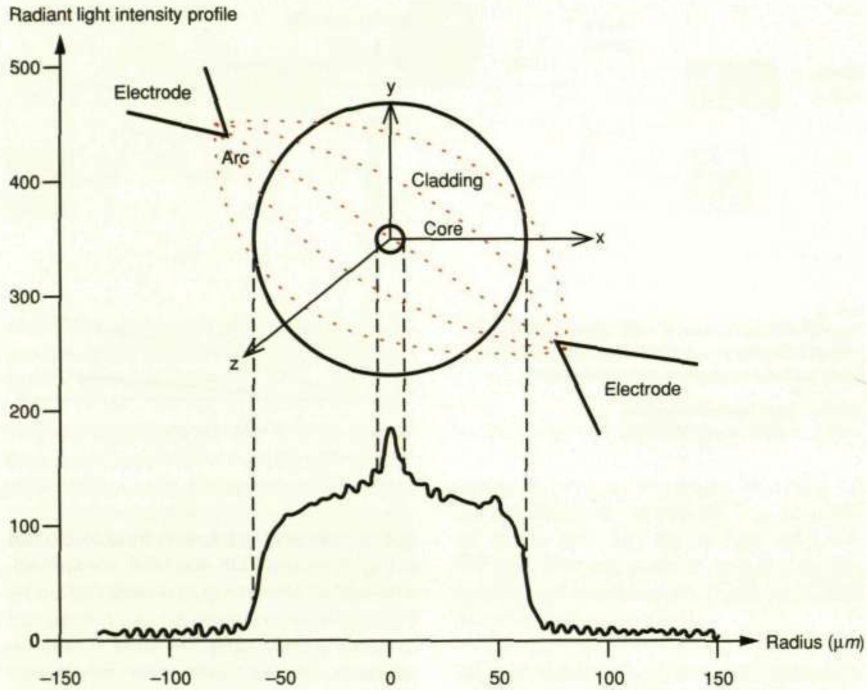


Fig. 5
Projection of the light-intensity profile of a heated fibre during electrical arc fusion. The intensity scale is in the grey level (maximum 255) of a digital image from a video camera

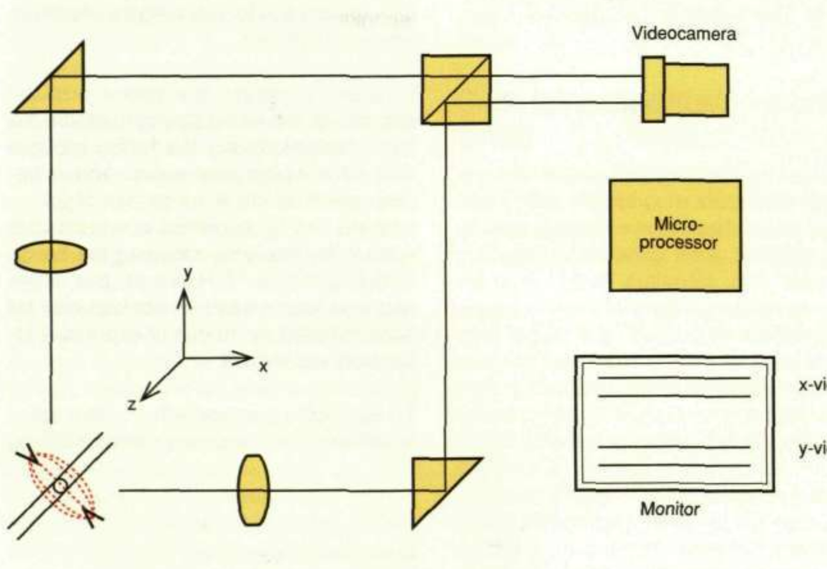


Fig. 6a
Schematic diagrams of the device for taking heated fibre images during splicing

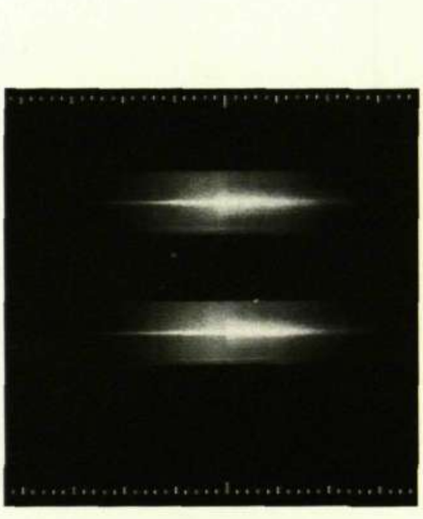


Fig. 6b
A heated splice image taken during arc fusion. An erbium-doped fibre (4 μm core diameter, left side) is spliced to a standard SMF (7.5 μm core diameter, right side)

of the heated fibre, as illustrated in Fig. 7. Furthermore, it is possible to calculate approximately the relative index difference Δ and the core diameter from the heated-fibre index profile during the fusion process. From this image-processing technique, we can understand better what really occurs when erbium fibre is spliced.[11]

Fig. 8 clearly shows what occurs during the fusion of an EDF/DSF splice. At the beginning of the fusion process, the diffusion speed of the EDF is much higher than that of the DSF, since the concentration of dopant is higher in the EDF (time A in Fig. 8). As the fusion process continues, the mismatch in refractive indices and

Fig.7
Comparison between the refractive index profiles of cold fibres and the heated fibre image profiles for

- (a) **Cladding depressed singlemode fibre**
- (b) **Erbium-doped fibre**
- (c) **Dispersion-shifted fibre made with the MCVD method**
- (d) **Dispersion-shifted fibre made with the OVD method**

Heated fibre index
Refractive index

core diameters becomes smaller and smaller, the change of the dopant concentration being much more significant than the change of the core size. The splice loss reaches a minimum point when the relative index differences almost match (time B in Fig. 8). Afterwards, the diffusion speeds of EDF and DSF are similar, but the maximum changing rate of the refractive index along the fiber axis is increasing rapidly at the edges of the heating zone. The increase of splice loss is mainly caused by these areas, not by the splice point (time C in Fig. 8).

Splice loss estimates

With the heated-fibre image processing technique the splicing process is no longer blind. The splice loss can be estimated from the measured core sizes and refractive indices. Since the core sizes and the indices are not uniform along the fibre axes in the fusion region, because of diffusion of the material, the loss calculation formula Eq.(1) derived from uniform MFDs for butt-joint splices[2] is no longer valid.

Instead of the conventional formula which considers only the mode field mismatch at the unfused splice point, the mode coupling theory should be applied when estimating splice loss.[8] In the splice loss calculation, the refractive index, the core centre position, and the core diameter should be considered as functions of the fibre axis z. The calculation of the refractive index fluctuation, i.e., the splice loss due to the

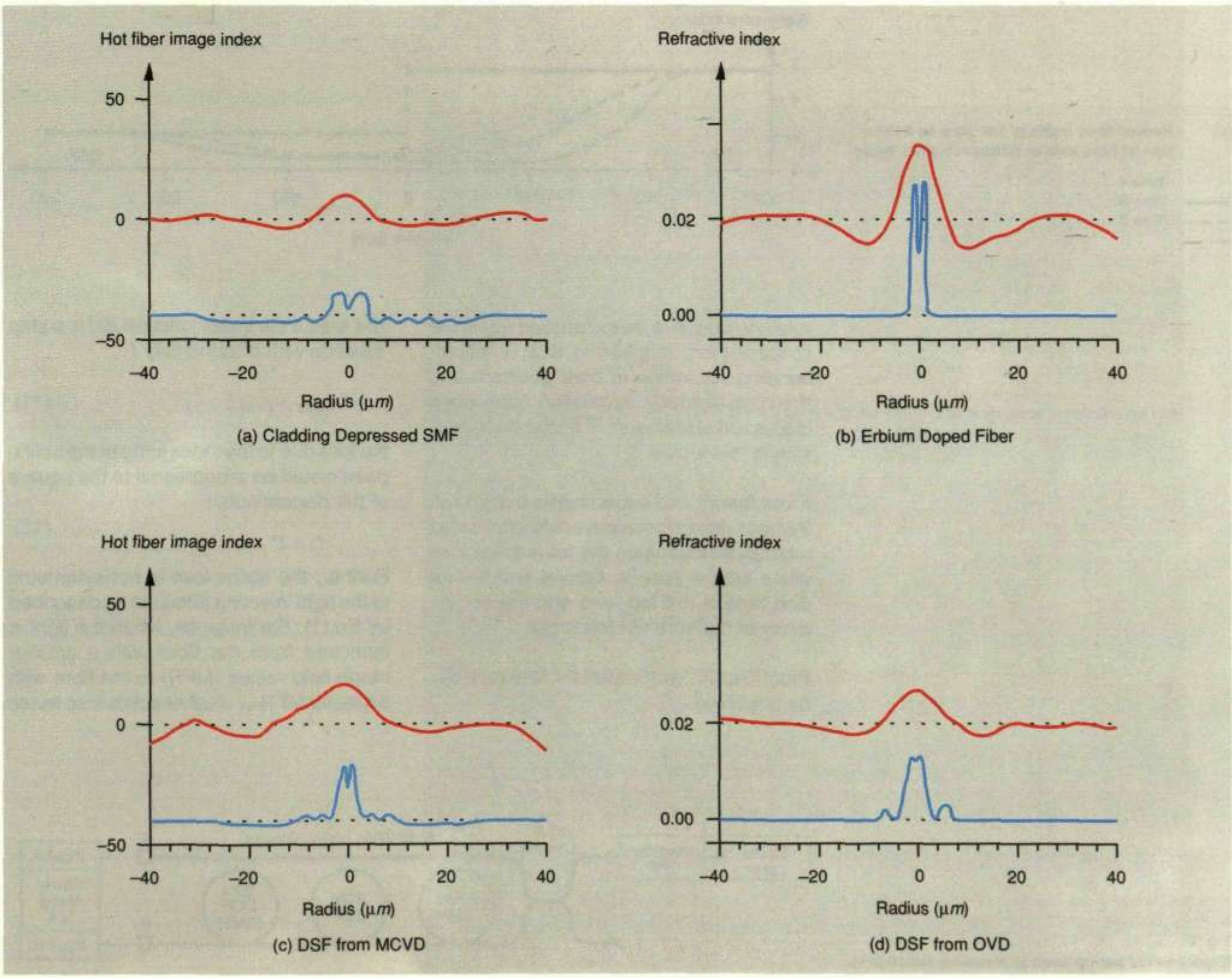


Fig. 8
Fibre parameters changing during arc fusion analysed with the image processing technique for heated fibre images

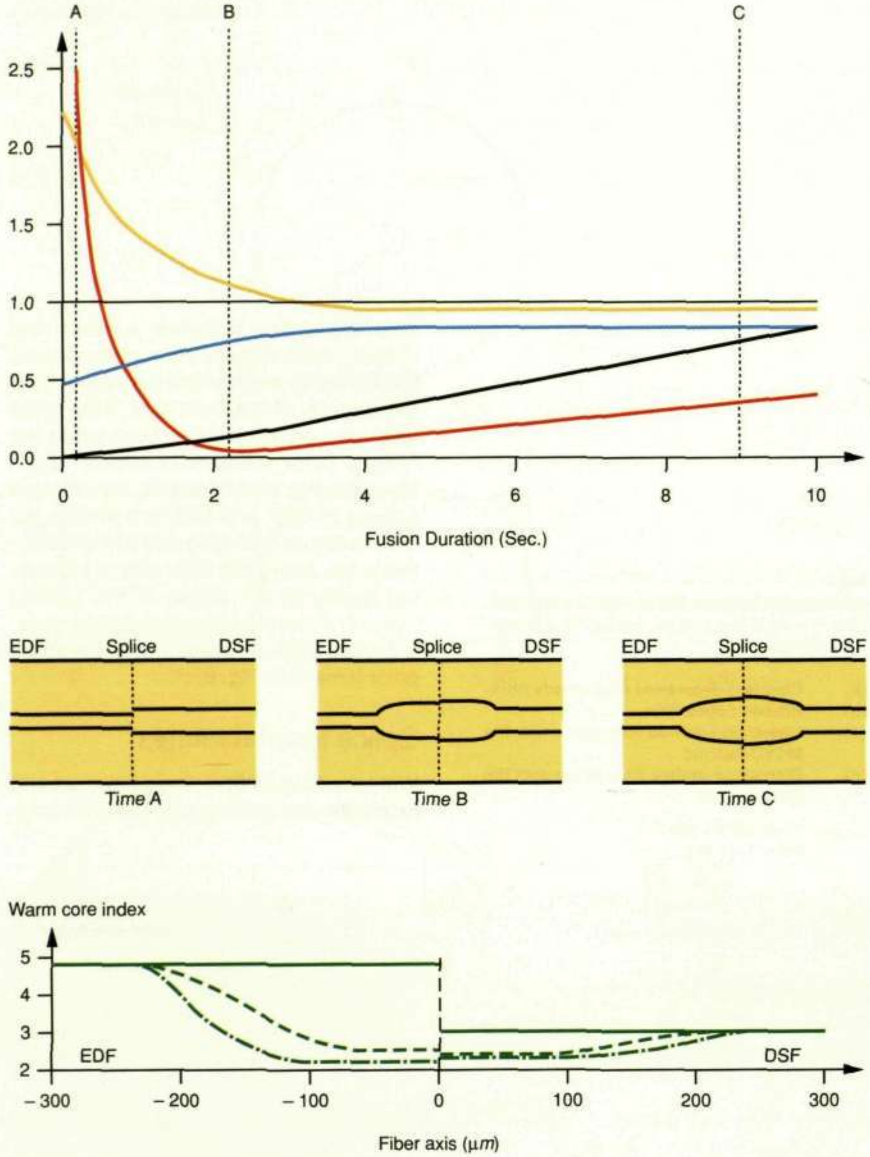
(a) ***Relation between splice loss and fibre parameters***

- Splice loss in dB
- Refractive index ratio: $\Delta_{EDF}/\Delta_{DSF}$
- Core diameter ratio: d_{EDF}/d_{DSF}
- Maximum index changing rate at edge of heat zone, unit 0.001 /µm

(b) ***Core size diffusion at different fusion times***

(c) ***Heated fibre index of the core as a function of fibre axis at different fusion times***

- Time A
- Time B
- Time C



relative refractive index function *i(z)* in the splice region, is given in Box B. Microbending (distortion in core position) and the core diameter fluctuation have been discussed elsewhere[8,12] and are not dealt with in this article.

From theory and experiments it is known that loss due to the relative refractive index mismatch dominates the total splice loss when erbium fibre is spliced and the fusion time is not too long and the eccentricity of the fibre not too great.

From Eq.(7), two important features can be observed.

The first, if the index function *i(z)* is a step function with discontinuity *I*

$$i(z) = I \cdot u(\frac{L}{2} - z) \tag{11}$$

the loss due to the index jump at the splice point would be proportional to the square of the discontinuity:

$$\Gamma_i \propto I^2 \tag{12}$$

Further, the splice loss is not symmetric to the light injecting direction as described by Eq.(1). For instance, when the light is launched from the fibre with a smaller mode field radius (MFR) to the fibre with a greater MFR ω_L, higher splice loss would

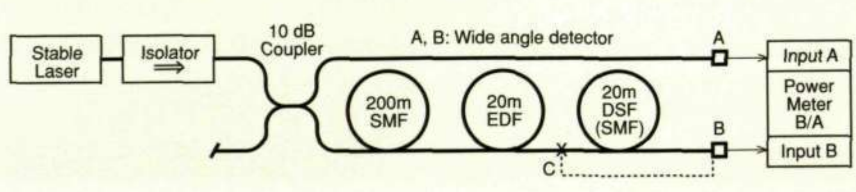


Fig. 9
Experimental set-up used to measure splice loss

be obtained from Eq.(7), and lower in the opposite direction for the same splice. If the splice loss with light transmission from smaller MFD fibre to larger MFD fibre is denoted $\mathrm{Loss}_{S\text{-}L}$, and that from larger to smaller is denoted $\mathrm{Loss}_{L\text{-}S}$, the relation between the losses measured in the opposite direction for the same splice can be approximately expressed by

$$\frac{\mathbf{Loss}_{S\rightarrow L}}{\mathbf{Loss}_{L\rightarrow S}} \propto \frac{w_L}{w_S} \tag{13}$$

in the ordinary long-distance communication frequency range. This counter-intuitive result is not predicted by the conventional butt-joint loss estimation theory Eq.(1), nor by the multimode fibre splicing experience. Fortunately, the relation given in Eq.(13) is supported by results from experiments reported in[7] (p. 154).

To verify the splice loss estimates, a set-up for splice loss measurement is established and shown in Fig. 9. In the set-up, an HP 81554SM high-stability laser source is used to launch a 1.55 µm signal through an isolator and into a coupler. From the coupler, a small portion of the signal is measured by detector A as a reference to avoid drifting. The other portion of the signal is filtered by 200 m SMF and 20 m EDF, respectively, to remove cladding modes as well as higher-order modes. The wide-angle detector B consists of an HP 81000BA bare fibre adapter, an HP 81050BL multimode fibre lens, and an HP 81521B optical head. The signals measured from ports A and B with two HP 81533A power meters are compared and displayed in the mainframe HP 8153A. Before splicing, end C of the EDF is measured at B (indicated by the dotted line in

Box B

We assume that the refractive index profile has the following form:

$$n(r,z) = n_0(r,z) + i(z)u(a-r) \tag{5}$$

where n_0 is the index profile before fusion. u(.) is a step function defined by

$$u(x) = \begin{cases} 0, & \text{if } x < 0; \\ 1, & \text{otherwise.} \end{cases} \tag{6}$$

From the mode coupling theory13, the splice loss due to the index changing Γi can be computed from

$$\Gamma_i = \frac{8Ck^2}{w^2b^2} \sum_s \left| \int_0^L i(z) e^{-i\beta z} dz \right|^2 \frac{\left[\int_0^a J_0\left(j_{os}\frac{r}{b}\right) e^{-(r/w)^2} r dr \right]^2}{J_1^2(j_{os})} \tag{7}$$

$$\beta = \frac{V^2 - 1 - \ln V + \left(j_{1s}\frac{a}{b}\right)^2}{a^2 \left[\sqrt{n_1^2k^2 - \frac{1}{a^2} - \frac{2}{w^2}} + \sqrt{n_2^2k^2 - \left(\frac{j_{1s}}{b}\right)^2} \right]} \tag{8}$$

$$V = n_1 ka\sqrt{2\Delta}\,.$$

Where a and b are radius of the core and cladding, respectively;

$k = 2\pi/\lambda$ is the wave-number and λ the wavelength;
C=10log e = 4.34;
J1(.) is the Bessel function of the first order;
j_{1s} is the s'th root of the Bessel function of the first order;
n_1 and n_2 are refractive indices of the core and cladding, respectively;
w is the mode field radius of the fibre from which light signal comes;

The total splice loss Γ is the sum of losses due to core microbending Γ_b, core diameter distortion Γ_c, and index variance Γ_i:

$$\Gamma = \Gamma_b + \Gamma_c + \Gamma_i \tag{10}$$

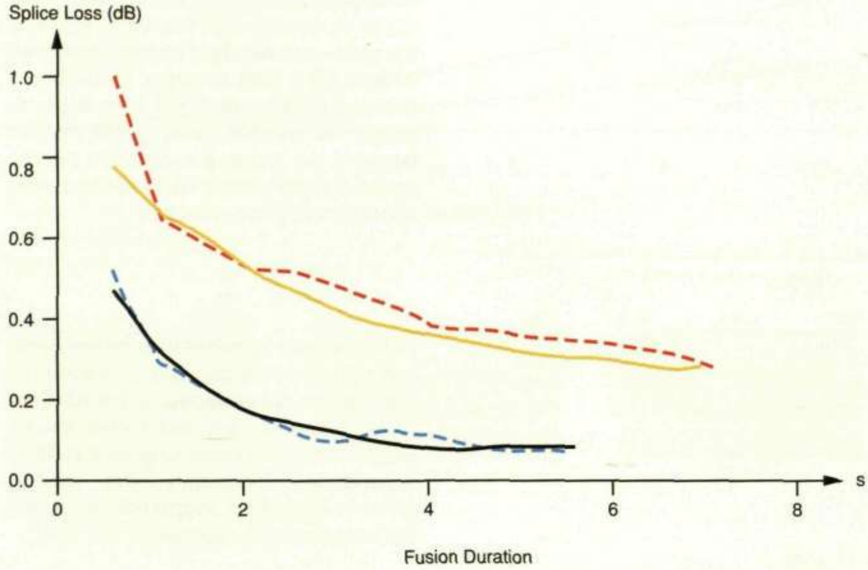


Fig. 10
Comparison between estimated and measured splice loss for fusion splices: erbium-doped and dispersion-shifted fibre at different fusion times

Estimated (Current 15.0 mA)
Measured (Current 15.0 mA)
Estimated (Current 16.3 mA)
Measured (Current 16.3 mA)

Fig. 9) as a reference. The mainframe display is then set to zero. After fusion splicing the EDF to 20 m SMF or DSF at C, the splice loss is measured and displayed at B. This set-up has a measurement accuracy of ±0.04 dB.[14]

In Fig. 10, two sets of measured splice loss are compared with the loss estimates of the fusion splices between EDF and DSF with a very small core-to-cladding concentricity error (less than 0.3 μm). The correlation between the estimate and the measurement is reasonably good.

Real-time control of fusion splicing

Since a good correlation between measured and estimated splice loss is obtained during the fusion process, real-time image processing can be employed for splice process monitoring instead of the active splice loss measurement.[15]. As a passive loss measurement, the image processing technique can also be used to control the fusion process in real time, as shown in Fig. 11.

When the fusion process has started, the splicer waits 0.6 second for the fibres to become overlapped and bright enough for the image to be taken. From the heated fibre image, the splicer calculates the splice loss due to the index mismatch. The fusion process will continue until the loss is smaller than a given threshold.

With the real-time image control technique, encouraging results are obtained for different fibre combinations. The measurement results for AT&T EDF to AT&T DSF, and for AT&T EDF to Corning DSF are illustrated in Figs. 12 and 13. In all the experiments, fibre ends were prepared with an acid stripping robot, splices

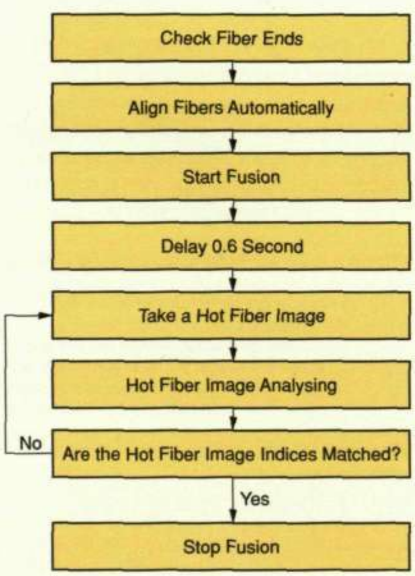
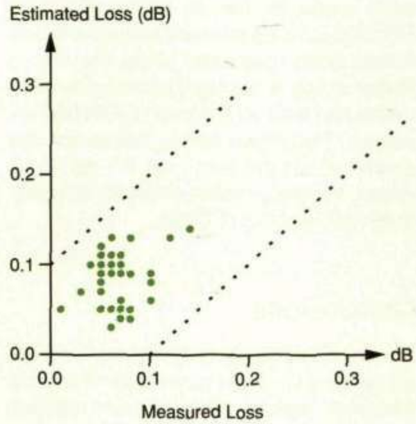


Fig. 11
Scheme of the real-time control process to erbium fibre splicing



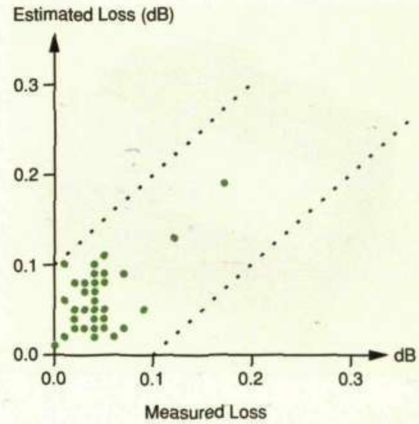


Fig. 12, left
Splice loss, strength, and correlation between estimated and measured values for 40 splices between AT&T EDF and AT&T DSF

40 Splices measured at λ=1.55 μm	
Mean loss	**0.07 dB (σ=0.02 dB)**
Mean strength	**276 kpsi (σ=44 kpsi)**
Yield (>200 kpsi)	**95%**

Fig. 13, right
Splice loss, strength, and correlation between estimated and measured values for 40 splices between AT&T EDF and Corning DSF40 Splices measured at λ=1.55 μm

Mean loss	**0.05 dB (σ=0.04 dB)**
Mean strength	**260 kpsi (σ=53 kpsi)**
Yield (>200 kpsi)	**90%**

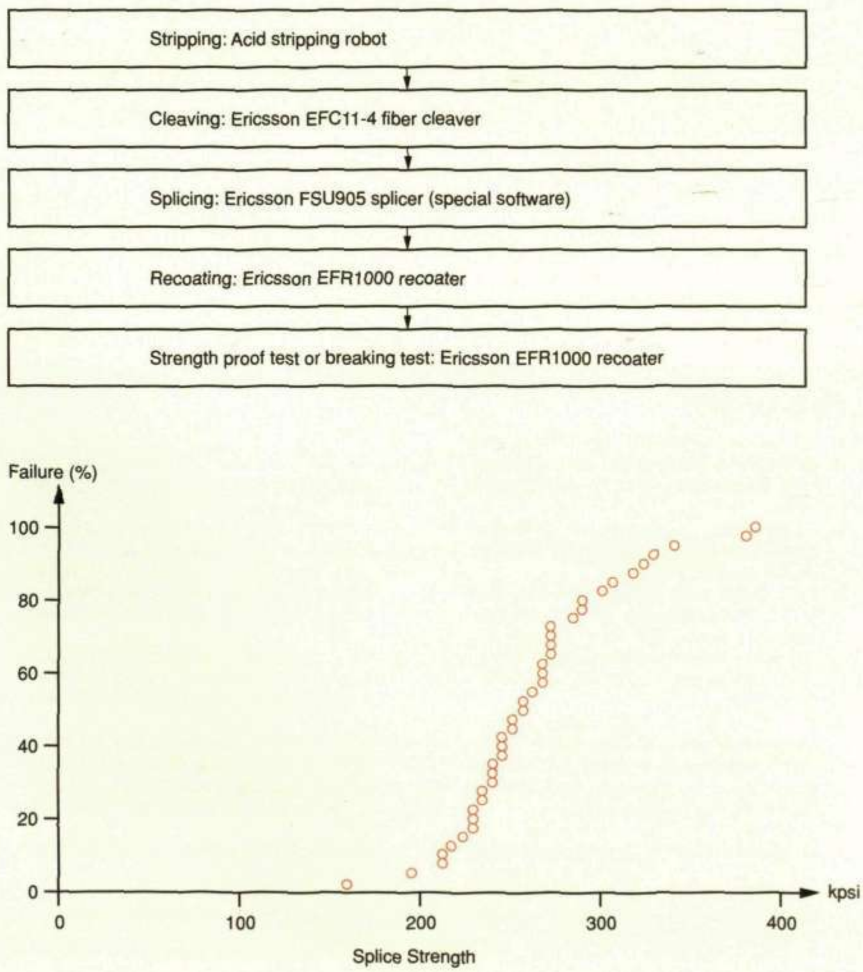


Fig. 14
Measured splice strength and the procedure used to make high-strength splices, AT&T EDF to AT&T DSF

Mean strength	**276 kpsi (σ=44 kpsi)**
Yield (>200 kpsi)	**95%**
Number of splices	**40**

Fig.15
Ericsson's FSU925 RTC fusion splicer with erbium fibre splicing software

were made by means of two Ericsson FSU905 splicers with new software, splice losses were measured using the set-up shown in Fig. 9, and splice strengths were measured with an Ericsson EFR1000 recoater. The mean splice losses for the combinations are less than 0.1 dB. In all cases, the mean splice strength is higher than 250 kpsi (≈ 1.7 GPa).

Conclusions

A unique technique is developed and implemented in a commercially available Ericsson optical fibre fusion splicer, Fig. 15, in order to splice erbium-doped fibres for fibre amplifiers.

During the arc fusion process, a significant amount of information about the core region of the splice can be inferred due to the high energy absorption and the corresponding thermal and luminescent emissions of the dopants in the core region relative to that in the surrounding cladding. The availability of emission data, coupled with the recently developed digital image processing technique, means that the refractive index and core diameter parameters can be dynamically determined. The mapping of the regions at – and immediately adjacent to – the splice allows accurate estimates of splice losses using mode coupling theory in contrast to less accurate loss predictions based on mode-field mismatch at the junction of the two fibres. The fusion process is automatically terminated when the heated fibre indices match.

Using this real-time image control technique, low splice loss, high strength, and a good correlation between splice loss estimates and measurements are consistently achieved.

References

1 R.J. Mears and S.R. Baker. *Erbium fiber amplifiers and lasers.* Optical and Quantum Electronics, vol.24, pp.517-538, 1992.
2 D. Marcuse. *Loss analysis of single-mode fiber splices.* Bell Syst. Tech. J. vol.56, pp.703-718, 1977.
3 M.P. Singh, J.O. Reese, T.M. Wei and D.G. Storch. *Low-loss fusion splicing of erbium-doped optical fibers for high performance fiber amplifiers.* IEEE Trans. on Components. Vol.13, No.4, pp.811-813, 1990.
4 H.Y. Tam. *Simple fusion splicing technique for reducing splicing loss between standard singlemode fibers and erbium-doped fiber.* Electronics Letters, Vol.27, No.17, pp.1597-1599, 1991.
5 T. Kashiwada, M. Shigematsu, T. Kougo, H. Kanamori, and M. Nishimura. *Erbium-doped fiber amplifier pumped at 1.48 μm with extremely high efficiency.* IEEE Photonics Technology Letters, Vol.3, No.8, pp.721-723, 1991.
6 S.G. Kosinski, K.W. Quoi, and J.T. Krause. *Low-loss arc fusion splicing of erbium-doped fibers with high strength for highly reliable optical amplifier systems.* Technical Digest of OFC'92, vol.5, p.231, San Jose, 1992.
7 C.M. Miller, S.C. Mettler and I.A. White. Optical *Fiber Splices and Connectors.* Marcel Dekker Inc, New York and Basel, 1986.
8 W. Zheng. *Loss estimation for fusion splices of singlemode fibers.* Proceedings of SPIE'91, Boston, USA, vol. 1580, pp. 380-390, 1991.
9 K. Shiraishi, Y. Aizawa, and S. Kawakami. *Beam expanding fiber using thermal diffusion of the dopant.* IEEE J. Lightwave Tech. Vol.8, No.8, pp.1151-1161, 1990.
10 W. Zheng. *Förfarande och anordning för att bestämma karaktäristiska egenskaper hos optiska fibrer* (in Swedish), Patent application no. 9201817-5, Sweden, 1992.
11 W. Zheng, S. Wong, O. Hulteén, S. Kosinski, and R. Rylander. *Erbium-doped fiber splicing with the real time control technique.* Technical Digest of OFC'93, San Jose, USA, pp 4-5, 1993
12 W. Zheng. *Real time control of arc fusion for optical fiber splicing* IEEE J. Lightwave Tech. Vol.11, 1993 (in press).
13 D. Marcuse. *Microdeformation losses of singlemode fibers.* Applied Optics, vol. 23, no. 7, pp 1082-1091, 1984.
14 J.D. Evankow. *Fusion-splice loss measurement procedure for dispersion-shifted fiber to erbium-doped fiber.* AT&T Engineering's Notes, private communication, 1991.
15 W. Zheng. *Förfarande och anordning för att skarva optiska fibrer av olika slag* (in Swedish), Patent application no. 9201818-3, Sweden, 1992.

ERICSSON

ISSN 0014-0171 Telefonaktiebolaget LM Ericsson 93301 Ljungföretagen, Örebro 1993

ERICSSON REVIEW

Radio Access Technology Evolution
Design for Electromagnetic Compatibility (EMC)
BusinessPhone 24 – a New Digital Key and Hybrid System
ACCIS – A New Tool for Processor Dimensioning for Call Capacity
ATM in Public Telecommunications Networks

3 1993

ERICSSON REVIEW

Number 3 · 1993 · Volume 70

Responsible publisher *Bo Hedfors*

Editor *Per-Olof Thyselius*

Editorial staff *Martti Viitaniemi*

Subscription *Eva Karlstein*

Subscription one year USD 30

Address S-126 25 Stockholm, Sweden

Fax +46 8 681 2710

Published in English with four issues per year

Contents

Cover
In usually crowded areas, such as Grand Central Station in New York – a cellular system with microcells is a prerequisite for extensive use of mobile phones

Radio Access Technology Evolution

Filip Lindell, Johan Sköld, Per Willars and Erik Nilsson

To meet the ever-increasing demands for greater capacity and new services, digital system technologies will have to evolve over the next decade. The future growth of cellular communications will be handled by the second generation of TDMA digital cellular systems that are now coming into service. During the next five to ten years, considerable capacity improvements and additional features will be introduced. When the cellular systems have reached the wideband service barrier, they will be superseded by the third-generation systems which will support those multi-media applications that have a need for variable bandwidth.

The authors describe the requirements imposed on the evolution of the digital cellular systems, discuss the suitability of CDMA and TDMA, and give some insights into the work on the next generation systems.

Cellular mobile telephony is now well into its second decade of unprecedented growth. The annual growth has been in the order of 40 percent, and there are no indications of saturation. In the middle of 1993, there were around 25 million subscribers in the world, and several market researchers predict 100 million by the year 2000. The fact that so many people will enjoy the benefits of mobile communications will have significant consequences not only for the way we spend our working hours and personal lives, but also – from a technical perspective – for the entire telecommunications network and especially for the radio technologies required to accommodate this tremendous growth.

Analogue cellular systems

There are a number of different analogue systems in use in the world today. More than half of the world's subscribers use mobile telephones according to the AMPS (Advanced Mobile Phone Service) standard. The two other prominent standards are NMT (Nordic Mobile Telephone) and TACS (Total Access Communication System). When these standards were developed, nobody could anticipate that cellular mobile telephony was going to be the success it has become today. The analogue standards were made for relatively few users who mostly had their telephones installed in cars. With the increased popularity of ever-smaller handheld telephones used in the street, inside buildings and in cars, it has become increasingly difficult to accommodate these users in the existing analogue systems.

The need for increased capacity has resulted in smaller and smaller cells being built. With the smaller cells comes the necessity to quickly and accurately hand over the call from one cell to the next before the user enters the third cell. As long as the cells were relatively large, the location and measurement processes could take

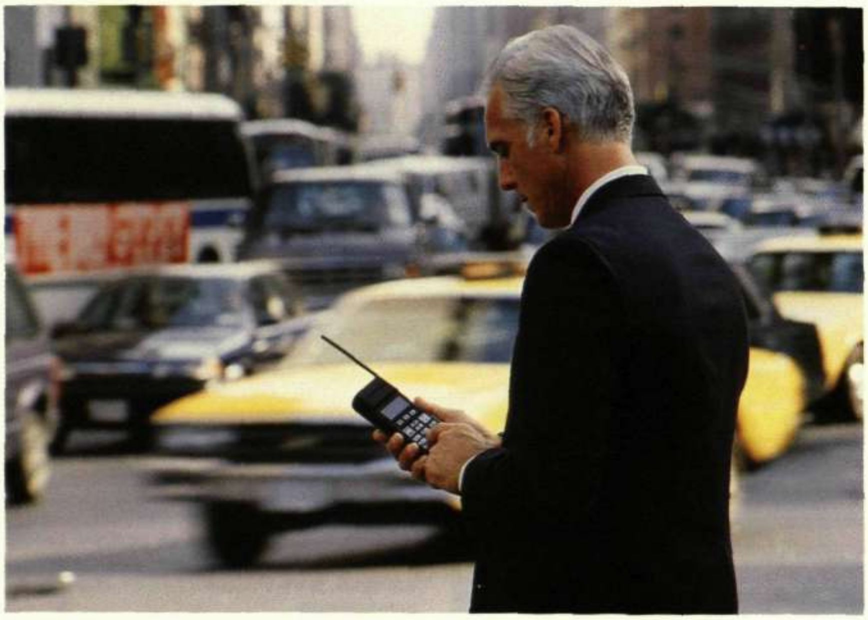


Fig. 1
Mobile telephones are a common sight among the crowds in the street almost all over the world




FILIP LINDELL
JOHAN SKÖLD
PER WILLARS
Ericsson Radio Systems AB,
Sweden
ERIK NILSSON
Ericsson Radio Systems Inc.
USA

several seconds, and the call would still be handed over to the correct cell, albeit with some delay. With the ever smaller cells being built today, the same processes are becoming too slow, which means that the call may not be handed over to the cell that provides the best speech quality.

The analogue cellular systems have been designed for voice communication. It is possible to transmit fax and data as well, by attaching modems at both ends of the connection. Practically, this has been difficult to handle, and non-voice applications constitute only a small portion of the total traffic today. With the advent of palm-top computers and miniature radio components, non-voice applications are expected to grow faster than voice during the next decade. The increased use of fax, data and short message services are placing new demands that will be difficult to meet with analogue technology.

Digital cellular systems

Several analogue cellular systems in use today meet with capacity constraints. They have served well for several years, and will continue to do so for many years to come. However, the future growth will be handled by the second generation systems – the digital systems. This is because these systems possess a number of features that make them more suitable for high capacity voice applications, and also for non-voice applications.

Firstly, they use the radio frequencies more efficiently. As can be seen in Table 1, the capacity increase compared with the analogue systems is 2 to 5 times from the outset. As digital systems evolve over the next few years, capacity increases of 10 to 20 times can be anticipated.

Box A gives an example of technical solutions under consideration that will increase the capacity of systems complying with the US digital standard to at least 10 times that of analogue systems.

Secondly, they have advanced functionality for small-cell and microcell environments. The much greater speed of handover decisions is made possible by the mobile-assisted handover function.[1] Handovers can now accurately be made at street corners where a user that turns a corner must immediately be handed over from one microcell to the next, Fig. 2.

In addition to higher capacity, cellular operators are interested in digital systems for the wider offering of services that can be provided for their subscribers. A promising future has been predicted for services like fax, data and short messages. Cellular systems have a great advantage in the amount of infrastructure deployed and the geographical coverage. With the move to microcells comes a decrease in output power and thus smaller subscriber units and longer battery life.

The industry expects great things from these services, as can be seen from the amount of devices and applications being proposed: PDAs (Personal Digital Assistants), wireless modems for notebooks, wireless E-mail, etc.

Box A
Capacity enhancement techniques under discussion for D-AMPS

Adaptive Power Control, APC	A scheme whereby mobile stations and base stations modify transmitted power levels in an effort to increase system capacity.
Antenna Diversity	A diversity technique whereby independent signals from two or more antennas are combined to improve the quality of the received signal.
Transmit Delay Diversity	A second delayed signal from the base station is transmitted from different antennas. The equaliser in the mobile station combines the two signals.
Digital Speech Interpolation, DSI	A scheme whereby the silence between speech bursts is used to carry other traffic.
Frequency Hopping, FH	A scheme which involves modulating data by a sequence of different carriers.
Discontinuous Transmission, DTX	A scheme whereby transmission power is reduced during periods of silence in order to minimise interference. Used together with FH.
Voice Activity Detection, VAD	Used together with DSI and DTX.
Coded Modulation	A technique for replacing channel coding, modulation and interleaving as currently specified in D-AMPS with combined schemes optimised for the application being supported.
Macro Diversity	Two or more base stations simultaneously transmit and receive signals from one mobile. Two techniques are used: selection diversity and simulcast.

Table 1

	AMPS	GSM	D-AMPS	PDC
Number of voice channels	312	~376	936	1137
Re-use plan	7/21	4/12	7/21	7/21
Channels per site	45	94	132	162
Erlang/km^2	3.5	8.8	13	17
Improvement factor	1	2.5	3.7	4.8

Digital cellular standards

There are three important digital standards in the world, one for each major economic region. The European standard is GSM, Global System for Mobile communications, the Japanese standard is PDC, Personal Digital Cellular. The North American standard for TDMA does not have an easily identifiable name. It has been given different names by different organisations. In this article, D-AMPS, Digital Advanced Mobile Phone Services, will be used.

Significant efforts have been put into the development and design of systems and terminals over the last years. The first releases are now in commercial service, providing a digital platform for further growth. This is important to remember, since the digital systems that are currently in service have implemented only a limited number of the capacity and quality features that will be provided mainly through software add-ons in the years to come.

Additional frequency bands

The analogue cellular systems are utilising frequencies in the 450 Mhz or 800-900 Mhz bands. The three digital systems are all in the 800-900 Mhz bands, with the important difference that the GSM and PDC operators were given previously unused frequencies, whereas the D-AMPS operators have to make a transition from analogue to digital of their often already congested radio frequencies.

Another difference is that the GSM and PDC standards also contain 1800 MHz and 1500 MHz versions, respectively. These higher frequencies have been licensed to new operators, to increase the availability of mobile communication services. At higher frequencies, the radio signal attenuation is greater which will reduce the size of the cell's coverage area at a given output power. The 1800 Mhz operators will then have to build a larger number of cells, which will drive the development of smaller and cheaper base station equipment.

In the US, parts of the 1800 MHz band will be vacated for mobile communications. They are currently used for fixed radio links, and it is not yet decided how the frequencies will be allocated and what technologies will be used. Many organisations favour a multitude of technical solutions. If this view persists, it will mean that operators can compete with technical differences, and not only with service and pricing as is the case in Europe and Japan.

Evolution of cellular into Personal Communication

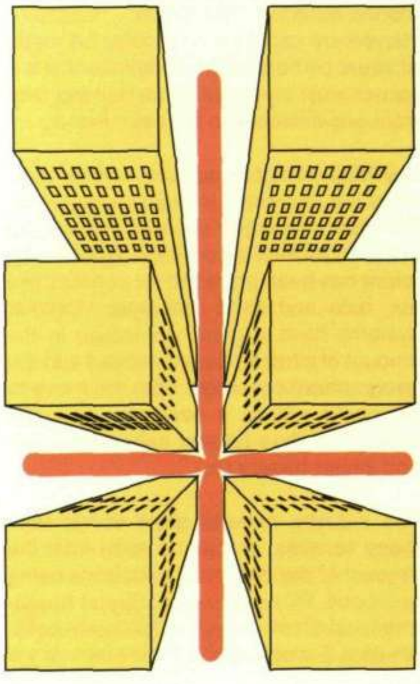
The broader meaning of Personal Communications is the concept of small pocket-size telephones that can be used everywhere at low cost. As new frequency bands are allocated and more operators are licensed, there will be an increasing number of systems, concepts and technologies under the umbrella of personal communications.

The term personal communications also implies a much greater number of users of mobile communications than today. 20 to 30 percent of the population in the developed parts of the world are expected to have "personal communicators" by the end of this decade. These will not only be for voice communication. A large growth is foreseen for various forms of text, data and electronic mail which will gradually increase the need for higher data rates in the systems. Speeds of up to 2 Mbits/s are being discussed, but such speeds will require a new generation of systems – the third generation.

The second generation will evolve to something close to the third generation (apart from the high data rate). The need for more capacity and the increasing number of operators competing for subscribers will place specific requirements on the evolution of second-generation systems. These evolved second-generation systems are sometimes referred to as generation 2.5. A graph of the mobile communication generations is depicted in Fig. 3.

With decreasing size and lower cost of both terminals and infrastructure equipment, cellular systems at 800 and 1800 MHz with their ubiquitous coverage are prominent parts of the family of personal communication systems. However, the increased competition from other personal communication systems and the dramatically increased use of mobile communications make a number of demands on the future cellular systems.

Fig. 2
A mobile telephone user that turns a street corner must immediately be handed over from one microcell to the next



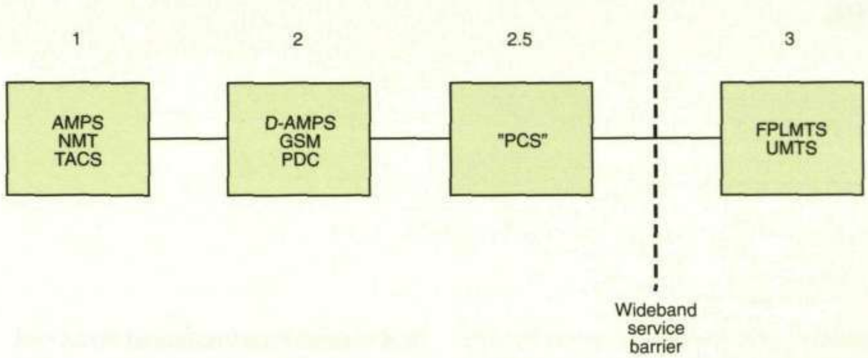


Fig. 3
Evolved cellular-based personal communications system, generation 2.5

AMPS	**Advanced Mobile Phone Service**
NMT	**Nordic Mobile Telephone**
TACS	**Total Access Communication System**
GSM	**Global System for Mobile communication**
D-AMPS	**Digital Advanced Mobile Phone Service**
PDC	**Personal Digital Cellular**
PCS	**Personal Communication System**
DECT	**Digital European Cellular Telephone**
FPLMTS	**Future Public Land Mobile Telecommunication System**
UMTS	**Universal Mobile Telecommunication System**

Demands on future systems and standards

Demands on both generation 2.5 and the following generation can be grouped into three main areas:
- Flexibility
- Quality
- Capacity.

Flexibility

For cellular-based personal communications, operational and service flexibility will be very important. Operational flexibility includes easy frequency/cell planning compared with today's systems. Also, the system must be capable of operating in different environments. The operator must have flexibility to allow different types of service at different cell types and to different subscribers, and maybe at different points in time. For instance, an indoor business system within company premises may have restricted access for users not employed by the company.

Quality

Quality includes speech quality, coverage and reliability of communications.

Good speech quality is very important. This must be considered when looking into speech codecs with lower bit rates than today, which are being proposed for capacity reasons. In fact, it is most probable that speech codecs with higher bit rates will be introduced to provide high-quality service.

Good coverage is needed to provide the services on demand. This will require the use of large cells for wide-area coverage, and specially designed smaller cells for local coverage of high-traffic spots, e.g. city centres, shopping areas, convention centres or parking garages.

Reliable communications in terms of a small number of dropped calls is desired.

Seamless handover is also required, i.e. the user should not be aware of the connection being transferred from cell to cell. This can be achieved by having a very fast synchronised, hard handover, or having soft handover, Box B.

Capacity

High capacity is also important for the next generation cellular system, of course. High cell capacity is achieved with efficient multiple access, coding and modulation schemes. Even more important, however, is high system capacity, i.e. the number of users that can be supported in a geographical area. For this, small microcells will be used where user density is high.

The cell capacity is constrained by the limited available spectrum and by co-channel interference. Thus, future systems must provide an efficient modulation technique as well as being resistant to interference.

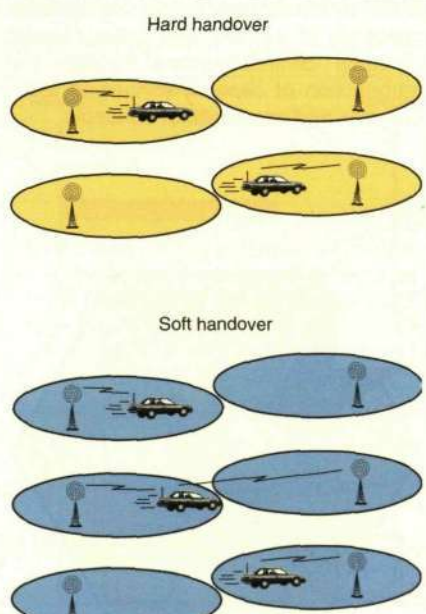
The system must be capable of adapting itself to changing traffic conditions. A soft capacity feature is desirable where a reduction of capacity in one area may be traded with higher than nominal capacity in another area.

Box B

The handover method most commonly used today can be described as hard handover, since the mobile station is ordered to switch its channel from the current base station to a new base station. The switch causes an interrupt of the speech transmission of several hundreds milliseconds – in some early analogue systems almost a second.

In a digital system, the interrupt can be reduced considerably if the base stations are synchronised. In this case, the mobile station will be capable of communicating immediately with the new base station without running through a synchronisation sequence. The interrupt in speech transmission will not be noticeable by the human ear.

The term soft handover refers to a situation when a mobile station in a border area communicates with two base stations simultaneously, i. e. macro diversity. As the mobile station moves from one base station area to another, it softly switches over to the new base station when this gives better transmission quality than the channel to the current base station. A prerequisite for soft handover is that the signals transmitted from the base stations are synchronised.



It is important that the quality is not compromised when maximising the capacity in future systems. There is always a trade-off between capacity and quality in a cellular system. An operator could use this to provide different quality to different subscribers and for different rates, or to let quality depend on the system load.

Hierarchical cell structures
To be able to meet the above-mentioned requirements, all cell types – e.g. large macrocells, street microcells and indoor picocells – must be allowed, Fig. 4. Furthermore, these different cell types must exist simultaneously in different cell layers, e.g. an umbrella cell covering a large number of microcells and picocells. Handover between different cell types must also be supported, of course. The ability to support these hierarchical cell structures is therefore one of the most important aspects of the access technique for personal communications.

TDMA evolution

The first digital cellular standards introduced in Europe, North America and Japan are all based on TDMA as access technology. As has been mentioned already, the first versions of the standards presently being deployed will provide 2 to 5 times the cell capacity of present analogue systems, thus meeting the needs of the growing cellular market in the first part of the 1990s. All three standards have the capability of evolving into cellular-based personal communications through the introduction of capacity-enhancing technologies and microcellular concepts.

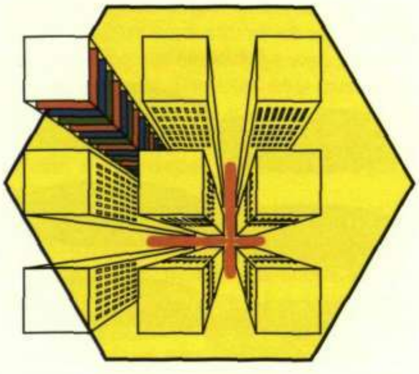


Fig. 4
An important requirement of a cellular communications system is the ability to build hierarchical cell structures

The D-AMPS standard is not yet fully digital. For compatibility reasons, the existing analogue control channel is used to access the system. The digital mobile telephone will then use a digital channel for voice or data traffic. A digital control channel is under development, however. It will allow for "sleep mode" in the mobile station, thus increasing the battery life; it will provide better control of hierarchical cell structures, support data services and SMS (Short Message Service), and have ISDN-based call control. The mobile stations may also be made fully digital, which will make them even smaller and cheaper.

The introduction of new technologies for TDMA, such as frequency hopping (FH) and adaptive channel allocation (ACA), gives increased capacity but also higher flexibility of the system by making frequency planning unnecessary.

Frequency hopping (FH-TDMA)
Frequency hopping is a way of increasing the system bandwidth of a cellular system, without changing the actual channel bandwidth. In GSM, slow frequency hopping is used. A set of frequencies is assigned to a group of users connected to a base station, and each user hops pseudo-randomly between the frequencies for every TDMA time slot. The hopping sequences are arranged within a base station such that no collisions occur, i.e. the users are orthogonal within the cell.

There are two capacity-increasing effects of FH. The first one is the frequency diversity. A narrowband system will experience deep fades of the signal level in frequency selective fading. This can be overcome by FH over a wider bandwidth giving frequency diversity. If information bits are coded and interleaved over several hops, the information can be restored.

The other effect is interference averaging. If the hopping patterns are random, the 'collisions' with users transmitting in other cells will also occur at random. If the number of frequencies used for hopping is large and the interleaving process is efficient, the collisions with strong interferers in surrounding cells will be randomised, resulting in an averaging of the interference from users and between users. This will increase system capacity, particularly in combination with discontinuous transmission, DTX, and power control.

The GSM recommendations specify both a random frequency hopping scheme, a speech coder with voice activity detection in combination with discontinuous transmission, and power control in mobile stations and base stations.

Adaptive Channel Allocation, ACA
The idea of ACA is to assign channels based on the instantaneous interference situation. For instance, a user close to a base station can re-use a channel used in the neighbouring cell, i.e. a re-use distance of one. A user at the cell border must be assigned a channel with a much larger re-use distance.

ACA is used in the Digital European Cordless Telephone (DECT) system with very promising results and has been proposed for D-AMPS[2]. With homogeneous traffic distribution, the capacity is increased by 30-40 percent. For inhomogeneous traffic distribution the gain is much higher. Furthermore, frequency planning is not required which is probably even more important since propagation prediction is very difficult when microcells are introduced.

CDMA with DS spreading

Whereas separation between users in FDMA and TDMA systems (Box C) is achieved through the use of different frequencies and time slots respectively, separation in a CDMA system is achieved through the use of different codes. Coding in CDMA systems is commonly referred to as spreading and may be performed either through FH or DS (Direct Sequence) spreading.

In DS-CDMA the energy of the user signals is distributed uniformly over the system bandwidth through the spreading process. The basic idea is that when several users are spread with different codes and added together, they will appear as noise-like interference to each other. In the de-spreading/decoding process, this noise will be suppressed by a factor called the processing gain.[3]

The addition and suppression of several interferers in DS-CDMA systems result in interference averaging. Thus, discontinuous transmission or variable-rate speech coding with power control can be used to increase capacity.

A requirement for DS-CDMA is that no interfering signal received is significantly stronger than the desired signal, since it would then jam the weaker signal (the near-far problem). Thus, a fast power control scheme must be used on the uplink.

Comparison between CDMA and TDMA

This section discusses some of the key characteristics and qualities of a TDMA system – evolved either with frequency

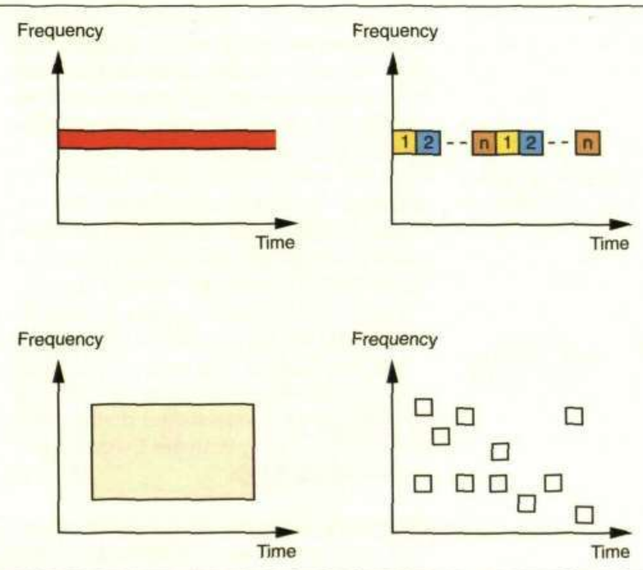
Box C

FDMA left
One channel occupies one frequency.

TDMA right
Example with n channels on one frequency divided into n individual time slots.

DS-CDMA left
Each user channel is spread over a broad radio channel. The individual user channels are distinguished by different codes

FH-TDMA right
Each burst of user information is allocated to a different frequency.



hopping as described in an article by B. Gudmundsson et. al.,[4] or adaptive channel allocation. For reasons of comparison, the discussion also looks into a DS-CDMA concept as described by K. S. Gilhousen et al.[3] The comparison is made for some of the key requirements of cellular personal communications: cell capacity, quality, operational flexibility and hierarchical cell structures.[10]

Cell capacity
Some conference papers[3] have claimed that CDMA is superior to existing FDMA and TDMA cellular systems, while others[4,5] have claimed that an evolved TDMA scheme will outperform CDMA.

One of the key properties for providing high cell capacity in DS-CDMA and FH-TDMA is interference averaging. It enables the system to be designed for the average instead of the worst interference case. It also allows full use of DTX and power control.

Orthogonality within a cell (e. g. a conversation does not interfere with, nor is it affected by, any other conversation in the cell) is a feature inherent in FH-TDMA. This will give an advantage over DS-CDMA, where most of the interference is generated within the cell.

Soft handover is more or less a necessity in FH-TDMA and DS-CDMA to reduce interference, although TDMA with ACA can achieve high capacity using regular hard handovers.

These issues have been debated frequently, and the debates will continue, but the overall conclusion is that evolved TDMA and DS-CDMA can provide comparable cell capacity.

Quality
There is no major difference between evolved TDMA and DS-CDMA air interface concepts in terms of speech quality. Instead, the speech quality is strongly dependent on the quality of the speech codec and the data rates used. The speech codec currently selected for the US DS-CDMA standard established during 1993 has a lower quality than the D-AMPS and GSM codecs.

Seamless handover will be achieved via soft handover or fast, accurate hard handovers, see Box B. Both types of handover require measurement reports from the mobile stations. An important tool for this is the MAHO (Mobile Assisted handover) function, which generates measurement results from the mobile stations on the quality of neighbouring cells.

Operational flexibility
To avoid the need for frequency planning, the new system should either plan itself – through ACA, for example – or have a reuse distance of one, which is achievable with both FH-TDMA and DS-CDMA.

In order to have the highest possible flexibility in the configuration of the system, it is desirable that the smallest spectrum unit allocated is not too large. This is a drawback of DS-CDMA systems, since the wide bandwidth of the radio frequency, RF, channel is a principle of its operation. In the proposal in reference[3], an RF channel bandwidth of 1.25 MHz is used.

Evolved TDMA can use a narrower bandwidth (200 kHz has been proposed) which will allow for flexible allocation in the frequency spectrum. Moreover, the FH system permits each channel to hop within the entire system bandwidth, thus maximising the frequency diversity. This may in particular be a useful feature in microcells, where the delay between multipaths is short and a very wide bandwidth is required to provide sufficient frequency diversity. Alternatively, the multipaths must be created artificially with multiple transmitter antennas using transmit delay diversity.

Hierarchical cell structures
The support of hierarchical cell structures is critical in cellular personal communications. This is needed for the operator who wants to provide different types of service in different areas. It is also needed to achieve high system capacity, by efficient employment of e.g. microcells. Different cell types could include macro/umbrella cells, microcells, indoor cells, cells for restricted access (e.g. in a company office), highway cells, etc.

The idea is that different cell layers/types should be made orthogonal to each other. The operator, not the access technique itself, must be able to set the rules of handover between cell layers. handover between two cell types should not be

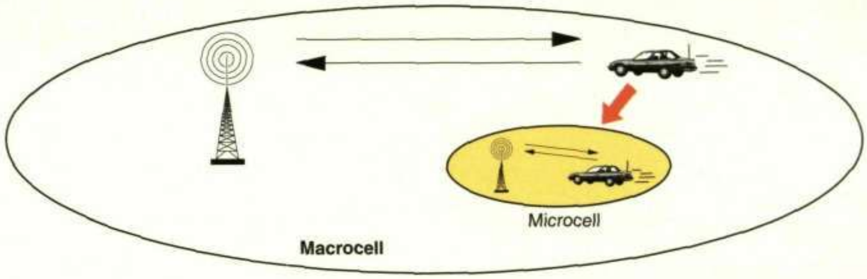


Fig. 5
When macrocells and microcells share the same frequency, the microcell will receive strong interference from a nearby mobile station in communication with a macro base station

forced by interference between them, since other criteria will be used for handover between cell types. These include mobile station speed, type of subscription and services, type of mobile terminal, etc. On the other hand, reliable handover must be supported for users who move between the different cell types.

The most commonly referenced overlay structure is a macro/umbrella cell overlaid on microcells. With this structure, fast users must be served by the macrocell, to avoid an extreme number of handovers. Slow users are allocated to microcells, to save capacity for the macrocell.

In principle, there are two ways of allocating spectrum for different cell layers: shared frequency allocation and separate frequence allocation. These are described below, followed by an investigation of methods of providing handover between RF channels.

Shared frequency allocation
Shared frequency is the method proposed for the DS-CDMA concept.[3] When microcells and macrocells share the same frequency, the microcell will receive strong interference from a nearby mobile station in communication with a macro base station, Fig. 5. A method called "desensitising" has been proposed as a solution to this near-far problem on the uplink. The output power of the mobile station in the microcell is increased so that the effect of the interference from mobile stations in the macrocell is reduced. Some of the drawbacks of this method are increased output power, resulting in reduced battery life, more interference in the system, and restrictions on the deployment of the microcells.

With shared frequency allocation, the cells are not overlaid but rather neighbours, i.e. handover is always required at the cell border. With a large number of microcells, this will lead to a large number of handovers for quickly moving vehicles, Fig. 6. This will place stringent requirements on the capacity of the radio access network. It is even doubtful that the radio interface will manage to perform the handovers with sufficient speed and reliability.

To conclude, the use of shared frequency allocation can to some extent be a solution for isolated microcells for local coverage. However, since it will make different cell types non-orthogonal, it cannot be used for a high-capacity, truly hierarchical cell structure.

Separate frequency allocation
Allocating different sets of RF channels to different cell layers gives the best flexibility for the configuration of cells in each layer. This is because the different cell types are now completely orthogonal and will not interfere with each other. Thus, the above-mentioned problems are avoided and microcells can be deployed independently of the macrocell structure.

The allocation of frequencies to the different cells layers could be either fixed or done in an adaptive way. Adaptive allocation means more efficient use of spectrum and easier planning.

Handover between RF channels
Since not all cell types can use one common RF channel, it is important to be capable of performing reliable handovers between RF channels. The mobile must make measurements on different frequencies for mobile-assisted handover (MAHO).

For TDMA systems, the idle slots can be used for MAHO measurements on other frequencies. The performance of MAHO

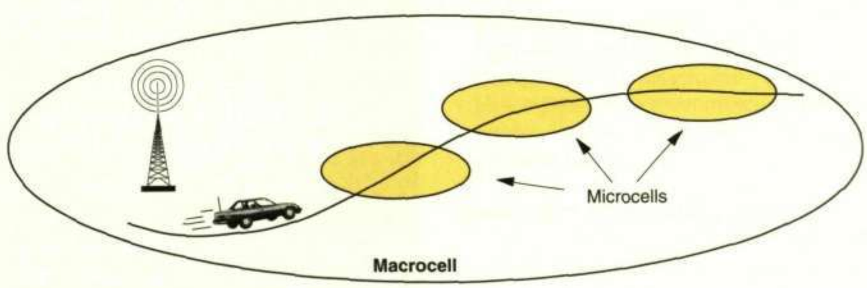


Fig.6
A mobile station moving at high speed would be better served by the macrocell. However, when macrocells and microcells share the same frequency, handovers are always required at the cell border

with ACA in a microcellular environment for an existing TDMA system has been evaluated.[7]

For a DS-CDMA system in general, the following methods of performing MAHO between RF channels have been discussed:

Dual Receivers
With an extra receiver in the mobile station, measurements can always be made on another frequency.

Downlink TDMA
By combining CDMA with a TDMA slot structure on the downlink, continuous reception is not required in the mobile station. The receiver can measure another frequency in the idle slot.

Short Measurement Slot
A short idle slot is introduced on the downlink, during which the mobile station must be capable of monitoring another frequency.

Discontinuous Transmission, DTX
DTX can be used when low bit-rate information is transmitted. A flag indicates to the mobile station whether the following frame/slot is idle, in which case the station can measure on other frequencies.

Pilot Beacons
Applicable to macrocell-to-microcell handovers only. Micro base stations transmit a pilot code on the macro RF,too; thus making the different cell layers non-orthogonal. The power level should be chosen such that the beacon pilot is strong enough to be detected at the microcell boundary, but weak enough in order not to interfere with the mobile stations connected to the macro base station. This will require careful power planning (together with detection thresholds, etc). In some situations this will work, in others it will not. Because of the near-far problem, a pilot beacon cannot be transmitted from the macro base station on the micro RF, since this would force the micro base stations to use the same output power as macro base stations. Thus, MAHO from micro to macro is not supported.

The third option seems most attractive to DS-CDMA in general, but its performance needs to be evaluated. The second and fourth options could degrade performance of fast closed-loop power control, if such a function is used in the system.

To summarise, handover between RF channels is more complex to achieve with a system based on DS-CDMA.

It should be noted that the DS-CDMA system proposed by K. S. Gilhousen et al[3] only supports the last option. This leads to serious constraints on the operators´ planning of a hierarchical cellular system.

Conclusion of CDMA/TDMA comparision
In the foregoing, systems based on two candidate access options for cellular personal communications have been studied. The first is TDMA evolved from today´s existing GSM and D-AMPS systems. The second is a stand-alone proposal based on DS-CDMA.

To evolve TDMA, there are two principal ways to go. One is to introduce frequency hopping, and the other is to use adaptive channel allocation.

The main improvement of these concepts for personal communications over current digital standards is the increase in cell capacity. This increase is of the same order for both schemes.

However, the suitability for personal communications depends not only on cell capacity. It depends heavily on the flexibility of the system. Normally, neither evolved TDMA nor DS-CDMA requires careful cell planning. Different frequency bands will be required for different cell types in order to get a truly hierarchical cell structure (and not only neighbouring cells). The support of hierarchical cell structures, without extreme planning, is easier to achieve with an evolved TDMA system than with the current DS-CDMA proposal, since the latter suffers some problems in supporting mobile-assisted handover between frequencies.

The choice of multiple access method for the third generation is still an open issue – none of the concepts discussed here is completely satisfactory.

The conclusion is that, for personal communication systems of generation 2.5, an evolved TDMA concept with FH or ACA is more suitable than the proposed DS-CDMA concept.

Third-generation requirements

The third-generation systems that will follow the currently emerging personal communication systems will be required to meet additional demands.

Besides the operational flexibility mentioned above, improved service flexibility will be needed. This includes substantial flexibility in the use of variable data rates. A wide range of data rates should be supported, from very low rate speech services – around 4 kbit/s – to high-rate data, up to 2 Mbit/s. This will require a variable bandwidth on the radio interface.

When these systems emerge, Asynchronous Transfer Mode, ATM, technology will be widely deployed in the fixed telecommunication networks. To make efficient use of the transmission flexibility in both the fixed network and the radio access network, a new generation of cellular ATM-compatible switches are required.

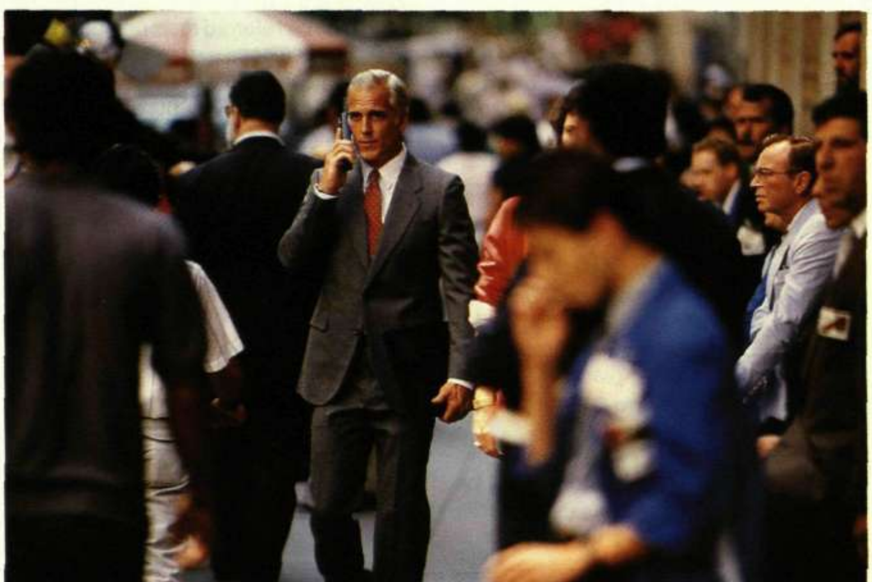
For third-generation systems, it is even more important than for second-generation systems to have an access method that supports hierarchical cell structures. Cells will be even smaller than today, but they will also be larger (e.g. satellite cells). All these different types of cells must be handled within the system, with increased quality.

High system capacity through high cell capacity and hierarchical cells will still be important. New techniques, using e.g. interference cancellation, that may increase cell capacity by an order of magnitude over current digital systems,[8,9] need to be evaluated.

The evolved TDMA and the CDMA concepts studied above do not fulfil all these requirements, e.g. regarding high data rates. Significant alterations are required to make them candidates for a third-generation system.

It is believed that third-generation technology will be based on either TDMA or CDMA, or a combination of the two. Much research is carried out in different parts of the world to determine a suitable concept. As an example of ongoing efforts, a European project is described below.

Fig. 7
Frequent use of mobile telephones in crowded areas will require a solution allowing for a mixture of cell sizes without need for detailed frequency planning



The CODIT project

CODIT, Code Division Testbed, is a research project into third-generation mobile communication systems.[11] Some ten companies are participating in the project, which is partially funded by the European Community in the RACE (Research and Development in Advanced Communication Technologies in Europe) program.

Requirements of the CODIT system concept

Some of the major requirements are:
- High-quality speech (same quality as wired line)
- Dynamically varying data rates up to 2 Mbit/s
- Video communication
- Terminal speeds up to 200 km/h
- Hierarchical cell structures
- Multi-operator environments (outdoor public, indoor office, residential base stations, etc).

The chosen technical approach

An advanced system concept based on DS-CDMA was established in the project during 1993. This concept will be refined and validated in the course of the project, and at the end of 1994 an updated concept with performance figures will be presented to standardisation bodies.

There is already an interim US standard based on the DS-CDMA concept.[3] However, this concept is aimed for a second-generation system and does not fulfil the requirements of a third-generation system.

The air interface for the CODIT project is focused on the flexibility requirements of a third-generation system, both with respect to operational demands and expected demands from a multitude of new, as yet unknown, services. The chosen DS-CDMA concept allows variable and mixed data rates. Also, by adjusting power and spreading ratio, services with different data rates can coexist in the same frequency band. A reduction in data rate directly leads to a reduction in the interference caused by one user, since the power can be lowered.

Three different bandwidths are specified for CODIT, but only the two lower ones will be demonstrated in a testbed.

The low-end bandwidth was chosen to be around 1 MHz. This is suitable for a flexible spectrum allocation to different cell layers in a hierarchial cell structure.

The high-end bandwidth was chosen to be around 20 MHz, due to the need to provide a high bit-rate service at reasonable complexity. It is likely that the high bandwidth can only be used in indoor picocells to limit interference. Therefore, an intermediate bandwidth is required to support relatively high bit-rate data services of 100-150 kbits/s in outdoor microcells. The intermediate bandwidth was chosen to be around 5 MHz.

Traffic channels are spread by one unique, very long PN sequence, and different users are separated by the phase of the sequence. The reason for this choice is primarily to avoid code planning in the system. A random phase can be chosen for each new call.

Macro diversity and soft handover will be used between adjacent cells in the same cell layer, but it will also be possible to handover calls to other cell layers that use different RF.

Tests and simulations

Three base stations will be used to test a hierarchical cell structure with a mixture of small and large cells in different cell configurations. Macro diversity and soft handover between base stations with the same frequency as well as seamless handover between cells with different frequencies will be essential functions to test.

Two mobile stations will be used – one of them in a van and the other on a trolley – for outdoor and indoor measurements, respectively.

Speech codecs with variable bit rates up to 16 kbit/s and data services up to 128 kbit/s will be used to test the performance of the test bed in different test environments.

The test bed will be built during 1993 and laboratory and field tests will be made during 1994. In parallel with these tests, the full system concept will be verified through computer simulations and analysis. The test bed implements a subset of the concept only, and the results from the test bed will be fed into the simulations to verify the results.

Summary

The second-generation, digital, cellular systems are evolving into generation 2.5, which may be referred to as personal communications. These systems must provide increased capacity, flexibility and quality. Especially the ability to build hierarchical cell structures is highly important.

It is shown that the current TDMA standards can be evolved to meet the demands on personal communications, whereas a proposed CDMA concept is less suitable in this regard.

The main addition when third-generation systems are introduced will be the great flexibility of user data rates, from low speed speech to 2 Mbit/s data. Whether these systems will be CDMA or TDMA – or both – is under study. Some information about the CODIT project has been given as an example of one ongoing study.

References

1. Eriksson, H. and Nilsson, P-O.: *Development of Radio Access in Cellular Systems - the Basis of Personal Telephony*. Ericsson Review *68* (1991):1, pp 2-9.
2. Andersson, H., Eriksson, H.,Fallgren, A. and Madfors, M.: *Adaptive Channel Allocation in TIA IS-54.System*,Proc. 42nd Vehic. Tech. Soc. Conf., Denver (1992), pp. 778-781.
3. Gilhousen, K. S.,Jacobs, I. M., Padovani, R., Viterbi, A. J., Weaver, L. A. Jr. and Wheatly, C. E.: *On the Capacity of a Cellular CDMA System*, IEEE Transactions on Vehicular Technology, Vol. 40, No 2, May 1991, pp. 303-312 .
4. Gudmundson, B., Sköld, J. and Ugland, J. K. :*A Comparison of CDMA and TDMA Systems*, Proc. 42nd Vehic. Tech. Soc. Conf., Denver (1992), pp. 732-735.
5. Meidan, R.: *Frequency Hopped CDMA and the GSM System*, Proc. Fifth Nordic Seminar on Digital Mobile Radio Communications, Helsinki (1992), pp. 343-345.
6. Shapira, J. and Padovani, R.: *Spatial Topology and Dynamics in CDMA Cellular Radio*, Proc. 42nd Vehic. Tech. Soc. Conf., Denver (1992), pp. 213-216.
7. Wallstedt, K., Almgren, M., Andersson, H. and Grimlund O.:*Micro Cellular Performance in a TDMA System*, Proc. 43rd Vehic. Tech. Soc. Conf., Secaucus, New Jersey (1993).
8. Teder, P. and Willars P.: *Capacity Simulations of a CDMA System with Interference Cancellation*, Proc. Fifth Nordic Seminar on Digital Mobile Radio Communications, Helsinki (1992), pp. 229-235.
9. Ewerbring, M., Gudmundson, B., Teder P. and Willars,,P.: *CDMA-IC: a proposal for Future High Capacity Digital Cellular Systems*, Proc. 43rd Vehic. Tech. Soc. Conf. Secaucus, New Jersey (1993).
10. Eriksson, H., Gudmundson, B., Sköld, J., Ugland, J. K. and Willars, P.:*Multiple Access Options for Cellular Based Personal Communications*. Proc. 43 Vehic. Tech. Soc. Conf. New Jersey (1993), pp 957-962.
11. Andermo, P. G. and Larsson, G.:*Code Division Testbed, CODIT*. Proc. 2nd. Int . Conf. on Universal Personal Communications, Ottawa (1993), to appear.

Design for Electromagnetic Compatibility (EMC)

Kjartan Tafjord

KJARTAN TAFJORD
Ericsson Telecom AB

An electronic product should have adequate internal disturbance margins, be immune to electromagnetic disturbances in its environment, and – inversely – should not emanate electromagnetic energy that might disturb other equipment. The generic term for these characteristics is ElectroMagnetic Compatibility (EMC).
The author describes the design principles and control tools used by Ericsson in order to ensure good EMC characteristics during the useful life of a product.

The ever-increasing use of electric and electronic equipment results in an increase in the number of disturbance-generating sources that may affect adjacent electronic equipment. As a consequence, stricter requirements have been formulated in order to limit the disturbance generated by different types of equipment and to strengthen the immunity of individual units to external disturbances. ElectroMagnetic Compatibility (EMC) is the term used to denote the requirements and characteristics of equipment as regards conducted and radiated emission as well as conducted and radiated immunity.

The EMC requirements to be met by a product are formulated by national laws and regulations, international standards and specific customer demands. The EMC characteristics of a product will depend on the physical implementation of the product; that is, its design.

The hardware in today's telecommunications systems is characterised by dense

A definition of EMC:

EMC, ElectroMagnetic Compatibility, is the ability of an equipment to function satisfactorily in its electromagnetic environment without introducing intolerable electromagnetic disturbances into other equipments in that environment.

To achieve EMC, a product must fulfil applicable requirements for conducted and radiated immunity as well as for conducted and radiated emission.

Essential basic standards for telecom products are CISPR 22 *(emission)* and IEC 801 *(immunity)*.

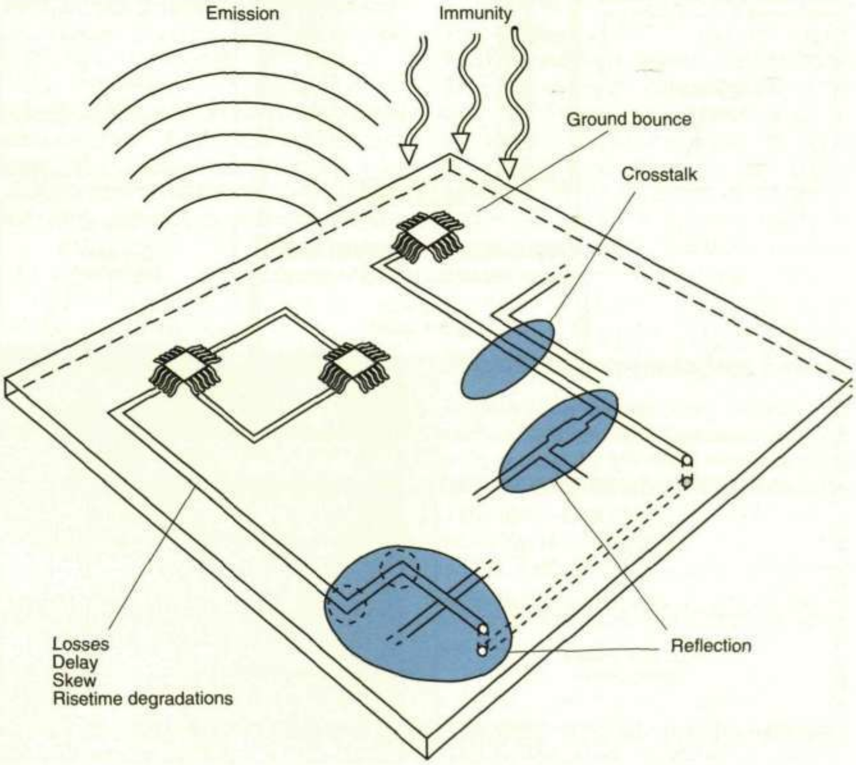
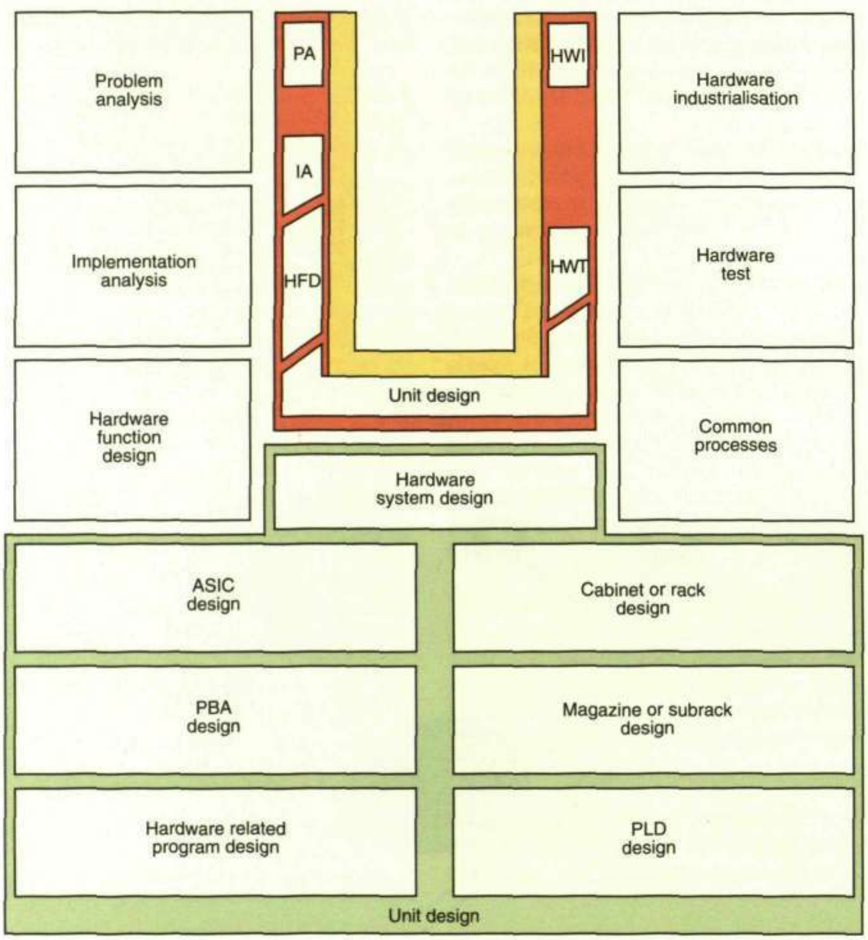


Fig. 1
EMC and signal transmission parameters on PBAs

packaging and high transmission speed. Dense packaging of high-speed signalling wires makes strict demands on the physical implementation of the product, if adequate EMC properties are to be ensured. This applies both to the packaging system and to the electronic design. When starting design work on a new product it is important to simulate to what extent the different levels in a given system structure will meet the specified EMC requirements. Shielding for EMC often affects the heat dissipation from a unit and, therefore, EMC and heat dissipation are often studied in parallel. Since crosstalk and reflection occurring in signal transmission affect both internal and external EMC characteristics, EMC and signal transmission requirements are coordinated when designing the wiring of a product.

Fig. 2
System design using HWDP



EMC requirements are often harder to meet than signal transmission requirements, and this is true of the layout of functions and components as well as of the wiring between components. Personnel with a good theoretical knowledge and long practical experience is a prerequisite for good results. Ericsson Telecom's hardware designers – i.e. designers of printed board assemblies, layouts and mechanical parts – receive basic training in EMC-oriented design.

The control tool in which the design for EMC is implemented is the well-documented HardWare Design Process (HWDP) shown in Fig. 2. In HWDP, the conditions of achieving good EMC characteristics in the product to be designed are formulated as requirements to be met from the start of design work up to the point where the characteristics of the designed product are verified. HWDP, which is available on the designers' workstations, is continuously improved in order to provide the best possible assistance in achieving adequate EMC characteristics. In the case of the packaging system, the control of EMC-oriented design work is part of another tool called the Electronic Packaging Design Project model (EPDP).

The documentation of the tools prescribes the design rules to be applied and what experts should be consulted in the different phases of the design work.

Partitioning of requirements to hardware level

External EMC requirements apply to the product as a whole – the system level – regardless of the physical structure or dimensions of the system. To be able to design a printed board assembly (PBA) that can be used in one or more systems, system requirements are partitioned to PBA level. Determining the requirements for immunity to external disturbances is the easiest part. System requirements are reduced by the internal system attenuation, i.e. the attenuation between the environment and the system level concerned. In the case of conducted disturbances, the disturbance levels specified in the requirements can be reduced by the attenuation obtained when decoupling cable shields, or by the attenuation in the filters used for

Fig. 3
Verification of radiated emission from an AXE 10 subsystem.
The picture shows the anechoic chamber with the equipment under test and the test antenna

unshielded cables. As far as radiated susceptibility is concerned, the target of study is the attenuation of the shields that separate the system level in question from the environment.

Radiated system emission at a given frequency amounts to the total radiation from the different radiation sources of the system – reduced, for each individual source, by the attenuation between the source and its environment. In a synchronous system with many equivalent coherent sources, the emission level will be $E = E_N + 20 \cdot {}^{10}\log N$, where N is the number of sources. If the signals are not phase-synchronised, the worst case condition will be $E = E_N + 10 \cdot {}^{10}\log N$.

A standard requirement value for radiated emission from a PBA would be approximately 40 – 50 dBμV/m at a distance of 10 metres, which requires installation of the PBA in a subrack having an effective shielding of 20 – 30 dB. (dBμV/m is the field strength relative to the 1 μV/m level.)

Building practice

All levels of the packaging system – from racks to subracks, connectors, cables, PBAs, printed boards and other components – must be designed so as to obtain good EMC characteristics without incurring large extra costs. Since conflicts may arise when determining radiation-reducing shielding and cooling, the treatment of these factors should always be coordinated.

The earthing of the packaging system should ensure the lowest possible impedance between the different parts of the earthing system. The impedance of AC signals in the earthing system is determined by the inductance of the different wiring paths. The self-inductance of a conductor that is long relative to its width is slightly more than 1 nH/mm. This means that the impedance is 2 ohms at a frequency where the length of the conductor is equivalent to one thousandth of the wavelength.

All metallic parts must be connected to the earthing system, and adequate galvanic contact must be provided for the bonding points of all metallic parts in racks and subracks; this requirement also applies to the mounting of subracks in a rack. The result is simplified design both of the electronic units to be contained in the packaging system and of the electrical wiring between the units.

If the best possible earth reference can be maintained at all levels of the packaging system, the effect caused by uncontrolla-

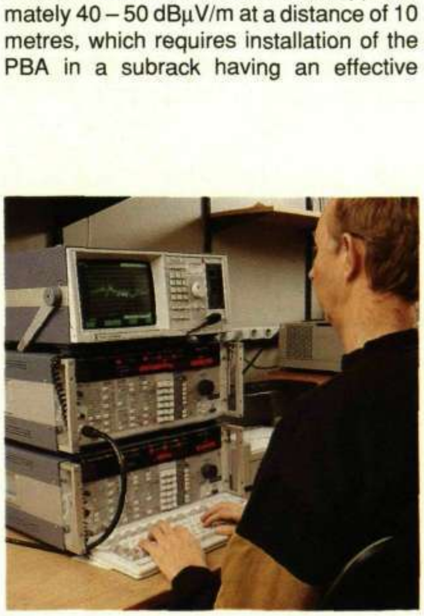


Fig.4
Verification of radiated emission from an AXE subsystem.
The operator and the test equipment are placed outside the anechoic chamber

ble signal currents in the earthing system will decrease. In digital systems, this means that the dimensioning of the density of the shields need only be adapted to the radiated emission caused by transversal currents in the signal wires. In the case of analogue parts with strict requirements for a disturbance margin, it is often the immunity requirements that determine the density of the shields.

Fig. 5 shows that signal currents may cause potential differences in the earthing system. The potential difference U drives a current I across the PBA to the connected cable. The PBA and the cable function as an antenna, which means that – to limit the outgoing radiation – the cable must be decoupled/connected with a very low impedance to the system shield. On the shield, the requirements for density are greatest at the cable connection points at subrack or rack level.

Figs. 6a and 6b illustrate transversal currents in signal wires and power supply. Together with the signal conductors, the current fed into the circuit forms current loops which act as loop antennas. The simplified expressions in Figs. 5 and 6a apply to transmitting antennas with small dimensions in relation to half a wavelength.

The shielding is distributed among the different levels of the packaging system, thus limiting the requirements for shielding effectiveness on the external shields to 20 – 30 dB. This keeps the radiated emission of the system at a low level and the immunity high even during maintenance work with open external shields. Besides, the maintenance cost of sealing openable shields can be kept low. Shields of this type are normally used at subrack level. Low shielding effectiveness in the external shield will necessitate more careful dimensioning of the internal wiring.

External cables connected to a subrack are designed either as shielded cables connected by shielded connectors, or as unshielded cables provided with filters. The connection field for external cables is included in the external shield, which also serves as earth reference. This part of the shield is where the requirements for density and low impedance are strictest.

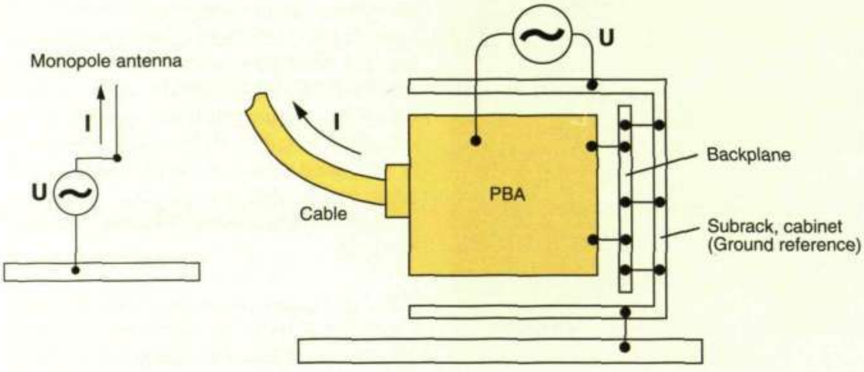
The internal wiring between component carriers is normally by means of connectors, the type of connector being determined by the signal density and the signal speed. For pin connectors, information is provided indicating the number of earth pins required and where to place them. In this way, the best possible EMC and signal transmission characteristics will be obtained at a given frequency.

The printed board assembly is the commonest type of component carrier. In order to provide at least one voltage plane and one earth plane, only multilayer boards are used. The earth wiring from subracks to printed board assemblies must be designed so that the radiation from the board and its cables is limited, Fig. 5. In order to prevent too small a distance between the signal conductors, the printed board technology must be adapted to the packaging density by means of a sufficient number of shielded conductor layers (Fig. 7, stripline). Besides, the characteristics of the earth and voltage planes must not be negatively affected by signal paths passing through the board and having too large insulating lands. The number of layers required is usually dependent on the requirements for limitation of internal disturbances to a level that ensures reliable peformance.

Fig. 5
Potential differences in the earthing structure cause longitudinal currents. Here the PBA and the cable act as an antenna that is equivalent to a monopole antenna

$E \leq (4\pi \cdot 10^{-7} \cdot f \cdot I \cdot s)/r$ V/m

(Dimensions, $<\lambda/4$)
s Dimensions in m
I Common mode current
r Distance in m



External cabling, connectors

The partitioned requirements concerning radiated emission from connected cables and connectors are stricter than the system requirements, because the radia-

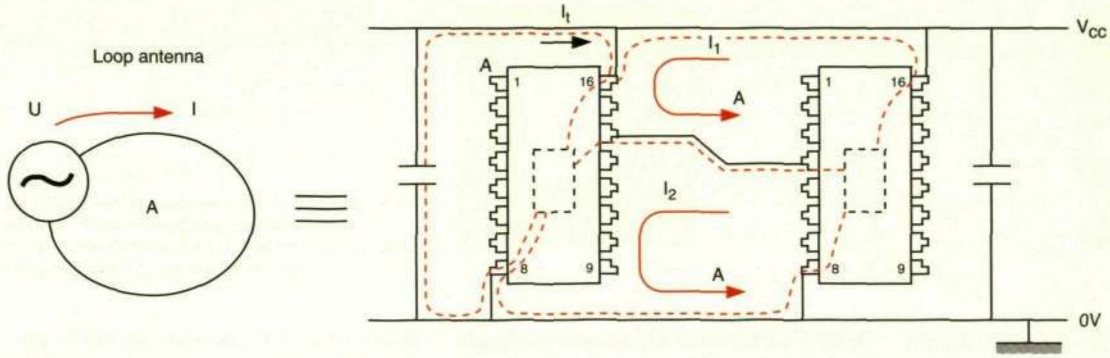
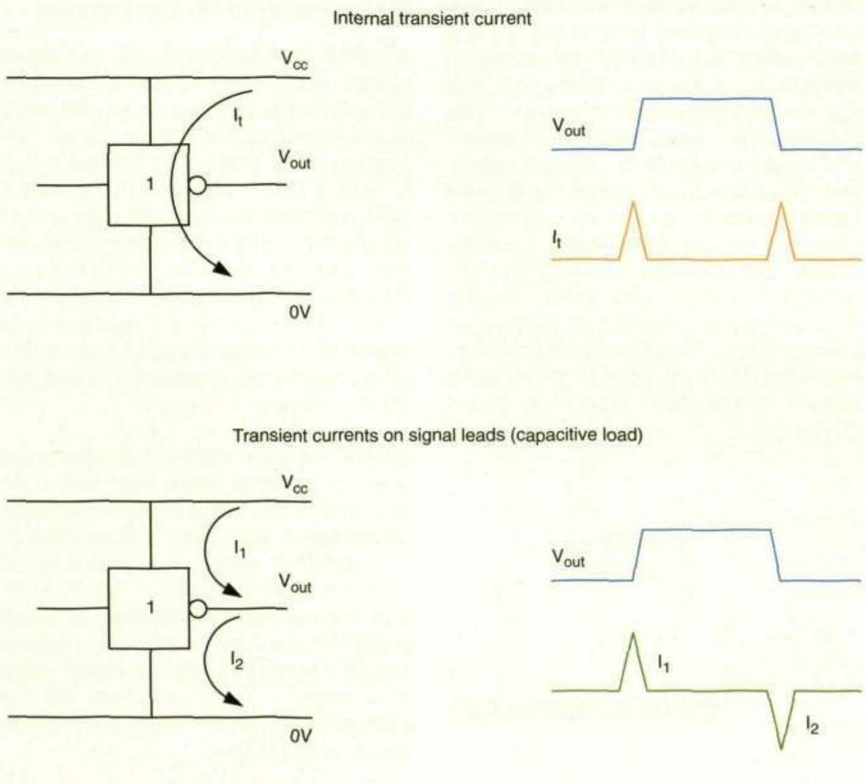


Fig. 6a
The power supply paths to circuits and the signal paths between the circuits together form current loops. These current loops act as antennas, whose equivalent is the loop antenna

$E \leq (132 \cdot 10^{-16} \cdot A\, I \cdot f^2)/r \quad V/m$

A   Loop area
I   Differential mode current (I_t, I_1, I_2)
r   Distance in m

Fig. 6b
The currents at the outputs from CMOS or TTL consist of two components: an internal transversal current that occurs on change of state, and the current fed to the signal wires



tion from different parts of the system is additive.

Cable shields and connectors must therefore be dimensioned so as to ensure a good margin to the partitioned system requirements, taking into account that aging may considerably deteriorate the characteristics of the shields.

External shielded cables are connected to shielded connectors, which in turn are connected directly to the earth reference of the connection field. Unshielded cables are filtered, and the filters are decoupled/connected to the earth reference of the connection field. The requirements for limitation of radiated emission from shielded cables and connectors are determined by the signals (including superimposed disturbances) sent over the transmission link: the greater the product of current times frequency, the stricter the requirement for shielding. Filters for unshielded cables are dimensioned so that the partitioned requirements concerning conducted emission, radiated emission and immunity are met. The immunity requirements are formulated so that, for example, induced conducted disturbances from wireless communications, electromagnetic discharges (ESD) and other transients will not affect the equipment.

Electronic design

The physical structure of the hardware is planned at an early stage of the design process. Before the structure is determined, the EMC characteristics of the proposed implementations are analysed in order to ensure that the EMC requirements will be met by the components of the packaging system. At this stage, designers decide what semiconductor technologies, package types, connector types and cables should be used. Either balanced or unbalanced transmission is chosen, and the rise and fall times required in signal transmission are studied. For digital products, wiring is dimensioned on the basis of the partitioned requirements concerning radiated emission. Crosstalk requirements decide the relative positions of the wire paths. The waveform of the signal on the inputs and the requirements with respect to radiated emission, decide how the signal paths should be terminated.

A typical connection runs from the output of a chip in a package on a PBA to the input of a chip in a package on another PBA. Series termination is an alternative in this case. A terminating resistor is placed so close to the chip output that no resonance can occur between the chip and the resistor, Fig. 8. The wiring on the printed boards (PBAs and backplanes) is designed for high-speed transmission, for example by using striplines. In these cases, the connectors between the boards must be either of the coaxial type, or pin or edge connectors with earth pins around the signal pins.

In analogue products requiring wide disturbance margins, the wiring is shielded to eliminate internal and external disturbances. In other words, the external immunity requirements are stricter as regards shielding than the external emission requirements.

Design of PBAs

The system requirements concerning radiated emission, partitioned to PBA level, decide what shielding requirements should apply to internal wiring on digital PBAs.

The wiring density and the shielding requirements decide what type of printed board should be chosen. However, in many cases the degree of internal disturbance decides how complex the structure of the printed board need be.

Depending on the shielding requirements, the internal wiring on printed boards and other component carriers is designed in either microstrip or stripline form, Fig. 7.

To achieve the best possible result, specialists in physical implementation of high-speed logic are responsible for locating functions and components on the printed boards. When the circuit diagram is loaded into the CAD system for pattern layout, the signal paths are roughly divided into signal categories, which are given denominations on the circuit diagram. For each signal category a description is made, showing what signal layers are allowed, and what rules apply to physical dimensions in order to ensure that the requirements concerning delays, reflection, crosstalk, etc. are met. Today, the CAD system cannot be used for simulating characteristics. Critical wire paths, usually clock paths, are simulated on the basis of assumed data. Where necessary, simulation with data from the actual layout can be used for verification during the layout phase.

Because of the radiated emission produced by resonances at harmonics of the clock frequency, even very short clock paths may require termination, Fig. 8.

Fig.8 is a simplified model of a chip-to-chip connection. R is a series terminating resistor which attenuates LC resonances.

Power distribution on PBAs

The PBAs are powered either by connecting 48 V to DC/DC converters on the PBAs, or by distributing the feed voltages from a common DC/DC converter, via backplanes, to the PBAs. In both cases, disturbances generated on the PBAs are prevented from spreading through the power connections by filters placed near the power connectors of the PBAs.

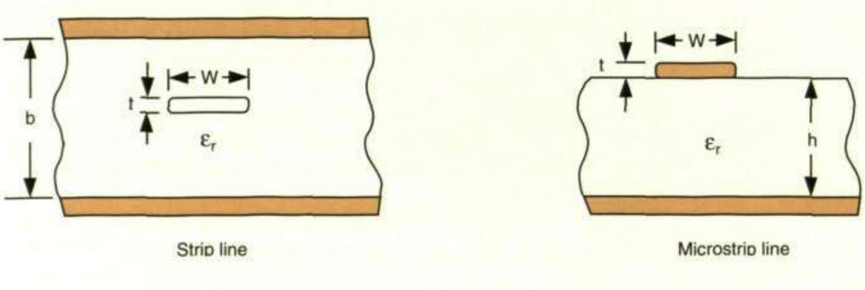
On the PBAs, logic voltages are often distributed via the voltage and earth planes. If these voltages are used for analogue functions, filtering near the current-consuming circuits is often required. Other voltages, for example 48 V, are in many cases distributed by means of wide conductors above the earth plane.

Decoupling of digital circuits

If a PBA uses both analogue and digital signals, or if it contains digital circuits operating at different voltage levels, the steep pulse edges of the digital signals will aggravate the problem of internal disturbances. It is then important to choose the right printed-board structure and the right decoupling method. A number of methods are used; for example, filter-separated local decoupling of circuits or functions, or – more often – common voltage and earth planes with strategically positioned decoupling capacitors of the same type and with the same value.

Two different decoupling methods can be used for common power feed from common earth and voltage planes. The method chosen in each case is dependent on the signalling speed, the printed-board structure, the risk of internal disturbances, and the number and location of signal paths through the earth and voltage planes. (A large number of signal paths may have a negative effect on the characteristics of the earth and voltage planes of the PBAs).

Fig. 7
The signal wiring on printed boards is in the form of either stripline wires or microstrip wires



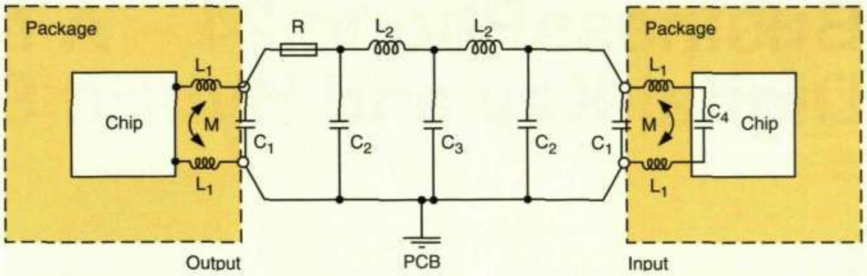


Fig. 8
Simplified circuit diagram for a chip-to-chip connection. R is a terminating resistor which will cause a low Q-value for LC resonances. Radiated emission and ringing will be reduced

According to the most common method for standard circuits, a sufficient number of identical decoupling capacitors are inserted between the earth and voltage planes. The capacitors are evenly distributed over the area of the PBA , without exceeding a specified distance from the packages. Ceramic chip capacitors with up to 1 mm long terminals are used. This is also the maximum length of the earth and voltage terminals of the circuits.

Another method uses a direct-connected chip capacitor at each earth and voltage connection point on the circuit; the capacitor is then connected to the earth and voltage planes. This method results in lower disturbance levels and may be suitable when decoupling circuits such as ASICs (Application Specific Integrated Circuits).

Type of capacitor and values are chosen taking into account the need for low impedance between the earth and voltage planes over a wide frequency band, and the need to eliminate resonances. Normally, a large capacitor is mounted on the DC/DC converter's output on the PBA for filtering and to ensure that the charge is sufficient to bridge momentary load variations. In parallel with the large capacitor, a sufficient number of identical ceramic capacitors are placed on the PBA. The inductance of the large capacitor generates resonance with the ceramic capacitors, but the resonant frequency is too low to be harmful. The inductance of the ceramic capacitors generates resonance with the board capacitance at a very high frequency where the Q-value is low.

Summary

Designers of products with good EMC characteristics must have a thorough knowledge and extensive experience. Besides, the EMC characteristics must be analysed when establishing the system structure, so that the right choices can be made at an early stage of the design work. Another prerequisite is anticipative planning in order to enable designers to determine what packaging system components will be required to meet internal and external EMC requirements in new products.

Designers need support in the form of control tools and rules, and to be able to provide correct input data for the design of PBAs they also need facilities for simulating signal transmission parameters on critical signal paths.

The tools available today do not permit reliable simulation of radiated emission from a PBA in its operating environment.

Ericsson Telecom has therefore chosen to train all its PBA designers in basic EMC-oriented design, to use specialists for analysing EMC characteristics during design work and to implement EMC control tools in HWDP and EPDP. The result is a marked improvement of the EMC characteristics of Ericsson Telecom AB's new products.

BusinessPhone 24 – A New Digital Key and Hybrid System

Bengt A. G. Andersson

In its basic version, Ericsson's new digital key system – BusinessPhone 24 –serves 16 extensions and 6 exchange lines and can be expanded to include another 8 extensions and 2 exchange lines. BusinessPhone 24, now introduced in some 35 countries, is a hybrid system permitting ordinary telephones to be used together with dedicated keysets.
The author describes the system and the services it provides.

BENGT A. G. ANDERSSON
Ericsson Business Network AB

BusinessPhone 24 is a new digital key system from Ericsson. The user's instrument can either be an ordinary telephone or a proprietary instrument (system telephone) providing enhanced facilities. BusinessPhone 24 has been sold in 31 countries, approved by the regulating authorities in the 17 countries where such an approval is required. The basic configuration is 16 extensions and 6 exchange lines contained in a flat-pack unit. By adding an expansion board, the system can be expanded to 24 extensions and 8 exchange lines.

Each digital extension line carries two 64 kbit/s digital channels which can be used simultaneously. With the aid of an analogue adapter, these two channels can be converted into two ordinary analogue extension lines. This increases the capacity to 47 accesses, one of which provides digital access for the operator. Four types of system telephones with varying characteristics are available.

Fig. 1
The central unit for BusinessPhone 24 and some of the instruments available. On top, the Executive model together with the Direct Station Select set with a button and a lamp for each extension and trunk. Below, the loudspeaking Standard model (left) and the Economy model (right)

Characteristics

BusinessPhone 24 is a digital system with an architecture similar to the ISDN 2B+D concept, providing two 64 kbit/s traffic-carrying channels and a separate 64 kbit/s signalling channel on each extension line. Development has been based on an investigation of identified and expected changes in customer requirements, technology trends, and ISDN penetration. The user interface for the different functions and cost penalties for added functionality have also been thoroughly studied so as to acquire the knowledge needed to design a cost-effective and easy-to-use key/-hybrid telephone system.

The result is a communication system well suited to today's need for communications in business and private environments and with a technology that makes the system adaptable to emerging changes in the user's business situation and telecom demands.

BusinessPhone 24

The basic system, Fig. 1, offers 16 extensions and 6 exchange lines built in a flat-pack unit. With an expansion board, the system is expanded to 24 extensions and 8 exchange lines. The system can connect two Direct Station Select instruments, DSS, for a receptionist or operator. Four different types of digital system telephones are available – Economy, Standard, Executive, and DSS – the characteristics of

Fig. 2
A modern key system does not improve your conversation but makes it easier for you to make your calls and provides the means to improve the response to people calling you

which are described below. Further ordinary analogue telephone sets and other analogue devices can be connected by means of analogue adaptors.

The analogue adapter splits the two digital 64 kbit/s channels into two separate analogue ports. A total of 23 of the 24 digital ports can be used, thereby increasing the number of extensions to 46. The remaining digital port is used by the system operator and for programming the exchange.

The system can be configured as a key system with all calls connected to all telephones, as a pure PABX system with incoming calls handled by the operator, or as a combination of these two alternatives. All facilities normally present in a key

Box A

Some important system features

Account code
If the user wants to charge the calling party (or a principal) for a call, he can enter an account code during conversation. The account code is then printed together with other relevant data in the Call Logging Information printout. This is an excellent feature for companies working in the service area, like travel agencies and firms of solicitors.

Class Of Service, COS
Eight classes of service are available. Each extension can have different COS for day and for night service.

Call Information Logging, CIL
Information about various phone calls in the system can be collected in CIL records. The record can be sent from the system through an RS-232C port to be printed or stored in a PC.

The details that are saved in the CIL record are:
- extension number of the recorded user
- trunk number
- digits dialled
- ringing duration before answer of an incoming trunk call
- date of call
- conversation start time
- conversation time
- account code entered.

Maximum 50 calls can be stored in the CIL buffer.

Direct Inward System Access, DISA
The Direct Inward System Access (DISA) facility allows a remote user to enter the system via a preprogrammed trunk (DISA trunk), gaining access to some of the system resources and facilities.

A call to a DISA trunk is answered with a special dial tone. For Direct-In-Dialling, the external party then dials an extension number and gets a connection to the wanted party. To get access to a trunk or to use abbreviated dialling, a code and a password are required. DISA makes it possible to make long-distance service calls, e.g. from home. The CIL record will contain the password used.

Trunk-Call Discrimination, TCD
Trunk-call discrimination is used to prevent extensions from making certain types of trunk call. The discrimination is based on day and night class of service and a table storing allowed digit combinations. For outgoing calls, the dialled digits are compared with the stored information. Up to ten digits may be checked and compared with up to 100 allowed digit combinations.

Night service
The system can be programmed to work in different service modes, day/night, with automatic or manual switching between the modes.

Music on hold
If an internal or external call is parked, the waiting party will hear music. The music comes either from the built-in music circuit or from an optional, external music source.

Call park in orbit
Any trunk call can be parked. Each extension has one orbit position in which to park a trunk call and make the call accessible from any extension. The paging function can be used to notify to wanted party to pick up the call.

External Call Diversion, ECD, to outgoing trunk
This function permits transit of external calls to a given external destination. Incoming calls to a certain trunk, the incoming ECD trunk, are diverted to another one, the outgoing ECD trunk. On the outgoing ECD trunk, the system dials the number programmed on abbreviated number 99. External call diversion can be programmed so as to be active during day-time, night-time or both.

PBX compatibility
BusinessPhone 24 can be connected to another PBX via trunks. These are assigned as PBX trunks and accessed by a programmed PBX-trunk access code.

Fig. 3
The Economy model has eight user-programmable keys in addition to the seven function keys. A built-in loudspeaker provides for handsfree monitoring and paging. Eight green-red light-emitting diodes present the busy/free status of available trunks

system of this size are available and certain new ones are added, Boxes A and B.

The two-wire connection of the digital telephone sets carries two connections, one voice and one data or two voice connections simultaneously. It also offers a separate signalling channel between the telephone and the exchange, providing for silent signalling on established calls. This makes it possible to reach a person who is already engaged in a voice connection over his handset, and announce another call through the loudspeaker of the telephone set. The user can accept or reject the new call by means of the menu buttons on his set. (Menu buttons are buttons used to select one of several alternatives presented on the display.)

Other valuable facilities include the possibility of locking and unlocking a telephone with a password, forwarding incoming external calls and the Direct Inward System Access, DISA, which makes it possible to access a station by post-selection and to make long-distance calls from an external telephone through the BusinessPhone system, using passwords for authorisation.

Extensions and exchange lines can be allocated to user groups, and call information can be logged, including call accounting data. The BusinessPhone 24 system can thus be used as a common communication system for several entities, so-called multi-tenant applications. The following external devices can be connected to the system:
- Battery for power backup. The battery is charged by the ordinary power unit, and no extra charger is needed
- Loudspeaker for public announcements and paging, e.g. in store rooms
- Source for music on hold and background music from the loudspeaker in the telephone when idle
- Outdoor type bell (loud bell)
- Printer or PC for call information logging
- IBM-compatible PC for system programming, connected either locally or remotely via a modem. The PC can also be used to read out and download customer data and also to read the logged call information.

Box B

Some important extension features

Paging
Paging can be performed from and to all extensions. The paged extensions hear a voice announcement in the loudspeakers of their phones. There are four types of paging, depending on whom the paging is intended for:

- voice announcement on a specified group of extensions
- voice announcement on all extensions
- voice announcement through a separate loudspeaker system
- voice announcement on all extensions and over the loudspeaker system.

Telephone lock
A telephone can be temporarily locked for outgoing calls. A password is used to release the lock.

Message waiting
When an internal call is not answered, or when the called party is busy, the calling party can leave a message. An extension can have several waiting messages, although only one from each source. The calling party can either use his own programmed message or one of the six messages common to all users.

Temporary change of Class Of Service
By means of his authority code, a user can activate - from any extension - the services permitted by his class of service. The authority code is active for one minute from dialling.

Conference
The conference facility allows four parties to be engaged in a conversation.

Background music
Each user can have background music through his loudspeaker, when his telephone is idle.

Automatic call-back on busy or no reply
In addition to automatic call-back on busy, BuinessPhone 24 offers call-back on no reply. When the called extension is activated - e.g. engaged in making a call - the system generates a call-back. If the call-back is not answered, a message will be left on the display.

Call diversion
The call diversion function contains a wide range of diversion possibilities:
- to a specified extension for all incoming calls
- to a specified extension when the call is not answered within a preset period of time
- to a temporarily defined destination
- to a specified extension when the original destination is idle
- to a specified extension when the original extension is busy.

A special dial tone reminds the user that call diversion is activated.

Camp on
Camp on busy can be used on busy extensions and busy trunks.

Last number redial and save dialled number
The last number dialled is stored and the number is redialled when the corresponding key is pressed. A dialled number can also be saved, and redialled, by means of the "Save dialled number" key.

Dual voice channel
All executive telephones are provided with a second voice channel. If an extension makes a call to an executive telephone whose first channel and handset are busy, the new call will be connected via the second channel to the microphone/loudspeaker in the executive set. The handset and the loudspeaking features can then be used simultaneously.

Information
The user can program his phone to present a message on the display of the calling telephone set. When this feature is activated, the message is displayed on the set. Six pre-programmed common messages and one message programmed by the user are available.

Reminder
A user can enter a time for alarm in his telephone set. At the given time, the phone gives an alert signal for one minute, or until the alarm is acknowledged. An alarm message is shown simultaneously on the display.

Fig. 4
The Standard phone has seven preprogrammed and 21 user-programmable keys. It is fully hands-free with built-in loudspeaker and microphone. Nine green-red diodes present trunk status

Telephones and terminals

Great emphasis has been placed on ease of use when designing the telephones and terminals. The buttons are arranged in three separate fields. First, the ordinary keypad, and below it a second field containing the buttons most frequently used, such as buttons for loudspeaking operation, hold, transfer/conference, program and clear function, and volume control. The third field at the side contains programmable buttons intended for functions that the user access frequently. The advanced model also has menu buttons which, together with the display, permit menu-driven operation.

Since the users will work within different trades and branches with varying needs in terms of features and packages of features, each customer can program his system and telephones to suit his special requirements. In BusinessPhone 24, each customer can choose his special package from the multi-feature menu and program the buttons on his telephones for best efficiency, both for business and private communications.

Maximum length of extension lines is 700 metres using 0.5 mm pair-cable.

In order to be able to distinguish whose telephone is ringing it is possible to choose between four different ringing characters. The ringing cadence is also different, depending on whether the call is an intercom call, a trunk call or an intercom callback.

All the telephone sets can be wall-mounted.

Economy model

This is the smallest of the BusinessPhone 24 system telephones. It is equipped with keypad, operational field with the most frequently used buttons, the same for all models, and eight programmable buttons with red-green light emitting diodes. It has a monitoring loudspeaker for on-hook dialling and paging.

Standard model

This model, Fig. 4, is a fully loudspeaking set with 21 programmable buttons, nine of which are associated with red-green light-emitting diodes for use as exchange line buttons. The large number of buttons enables the user to create an individual set of easy features, accessible through a touch of a button only.

Executive model

This is the most powerful model, loudspeaking and providing access to all system features, Fig. 5. It has a 2 x 16 character display which permits menu-driven operation by means of three menu buttons, each of which is clearly related to the appropriate choice. The display is also used for text messaging and information. The Executive set comes with 21 programmable buttons; the nine lowest ones have green-red LEDs for supervision of the exchange line traffic.

An advanced version of the Executive model is equipped with a built-in modem with a V.24 data terminal connector. This version must be used when the data communication functionality of the exchange is utilised.

Direct station Select, DSS

This unit has 36 buttons, each equipped with a red-green light-emitting diode for supervision of extensions and exchange lines. The capability to call an extension by simply pushing one button makes it a useful tool for operators and receptionists, Fig.6. The light-emitting diode at each button continuously presents the state – busy or free – of associated extensions and trunks. If the exchange is built out to more than 24 extensions by means of analogue

Fig. 5
***The Executive phone has a display with 2 x 16* characters for text messaging. The user can separately program the keys in the three left-hand rows for direct access to internal or external numbers or for other functions. For operators, the DSS unit with its 36 buttons and lamps showing the busy/non-busy state for extensions and trunks is a valuable tool**

Fig. 6
The Direct Station Select (DSS) set has 24 lamps that present status of extensions and eight lamps that present status of trunks. Trunks and extensions are accessed by the keys beside each lamp. The DSS is used as an add-on to a telephone

adapters, the DSS will work in two planes, supervising all extensions. If the system is configured for two operators, one of the DSS functions performs shifting between the two.

Analogue adapter
The analogue adapter, Fig. 7, converts the digital 2B+D interface of the extension line into two ordinary analogue telephone line interfaces. Like other terminals it can be located up to 700 metres from the exchange, allowing an additional 1300 ohm line to the analogue telephones or other analogue devices.

Features

The system supports all features normally found in key systems of this size, both from the system's and a user's point of view. All users can program their telephones so that any feature can be activated through the touch of one button. Some important features are shortly described in Boxes A and B.

Maintenance and service

The system is delivered with system and customer data set to the values most frequently used. Reprogramming on site can be executed from any extension equipped with the Executive telephone or from a PC connected to the exchange, locally or via a modem.

In BusinessPhone 24, any exchange line can be used for connection of a remote PC using the Direct Inward System Access, DISA, function. In this way the system can be programmed and customer data read out and stored at a maintenance centre and, when required, downloaded to the exchange.

The exchange is contained in a small flat pack which is replaced when faulty. Advanced internal fault diagnostics is therefore not required.

Technology

BusinessPhone 24 is a digital telephone system using standard PCM technology *for transmission and switching*. The design is built around three different ASIC chips supporting a 2B+D technology with a proprietary protocol for information exchange. This makes it possible to use a two-wire connection between the central unit and the telephones and other terminals. The system is controlled by device processors supported and supervised by a central processor. This concept has resulted in a compact and cost-effective design with a flexible system architecture, easy to adapt to future requirements. For technical data, see Box C.

The system is non-blocking and contains a number of switching links with sufficient capacity to handle all calls that can be requested simultaneously.

The memory has a seven-day memory backup. A feature that automatically connects one or two trunks (depending on version) to ordinary telephones in case of a power failure is also available.

Design

BusinessPhone 24 consists of four main printed board assemblies built into a flatpack box. The following boards are included:
- Common control board
- Board for sixteen extensions and power supply
- Board for six exchange lines
- Expansion board for eight extensions and two exchange lines.

BusinessPhone 24 is designed according to applicable European power safety regulations (IEC 950) and EMC requirements (CISPR 22, class B).

Fig. 7
The Analogue Adapter is used to split the two channels of a digital extension line into two analogue lines for connection of ordinary phones or other analogue devices

Documentation and sales support

Products like BusinessPhone 24 consist not only of equipment but also of descriptions, rules, advice and other documentation needed to support the operation of the equipment and the users' need for efficient and continuous communication services during the life of the exchange. These items have been created and developed along with the design of the system. The documentation is structured according to Ericsson standard.

Two binders are supplied to the sales and support agent. One of these binders contains the hardware and software documentation required for evaluation purposes; the other contains all information necessary for the operation and maintenance of the system. Installation and maintenance instructions, presented so as to be easy to read, understand and apply, are to be found in a booklet which accompanies each central unit, together with two sets of User Guides. A Quick Reference Card comes with each telephone set.

In addition to the documentation delivered to customers and users, various sales support material has been produced to depict the system, to illustrate the service it supports, and the benefits gained by the user. The material available includes brochures, data sheets, posters and advertising matter.

Conclusion

With Businessphone 24, Ericsson has a new system in the range from 4-6 to 24 extension lines in its product portfolio. The new product is compact, has all the features required for a modern key system and includes a maintenance system designed to give the customer full support and safety for his telecommunication facilities.

The most important new feature to be introduced in the system is adaptation to public ISDN. Investigations are also in progress – with valuable assistance from Ericsson's sales organisations and customers – to create a base for further development of the system.

Box C

Technical Data

Extension line data	
Maximum length,	
system phone	700 m
DTMF phone	*1300 ohms (including* phone resistance)
Trunk line data	
Loop resistance	2,200 ohms
Digit transmission	DTMF according to CCITT Q.23
Decadic 10 Hz	Ratio 40/60 (or, on market request, 33/67)
Transmission data	
Crosstalk attenuation	75 dB
Capacity	
Extensions	Up to 24 (basic version, 16)
Trunk lines	Up to 8 (basic version, 6)
Dimensions and weight.	
Wall cabinet, incl. power supply	
Height	432 mm
Width	290 mm
Depth	90 mm
Weight	7 kg
Power supply	
220/230 VA, 110/127 VA, 47 - 63 Hz	
Power backup with battery	
Power consumption	
Idle: 40 VA; Busy: 90 VA	
Music input	< 2 000 ohms
Environment	
Temperature	+5 °C to +40 °C
Relative humidity	5 – 90 % non condensing

ACCIS – A New Tool for Processor Dimensioning for Call Capacity

Magnus Höglund and Björn Kihlblom

Enhanced functionality requires more processing power. A number of control systems – APZ – with different processors are available to AXE 10, to meet the wide range of customer needs. To be able to tailor APZ to the needs of each individual exchange, a dimensioning system called ACCIS – AXE 10 Call Capacity Information System – has been developed.

The authors describe how the central processor of the control system in AXE 10 is dimensioned with respect to call capacity.

The call handling capacity of a telecom exchange is determined by the processor capacity, the load per call for the processor for different types of call, the traffic mix in the node, and the percentage of the processor capacity available for traffic handling.

When new functions or new lines are added to a telecom exchange, most of the factors just mentioned will change and the call handling capacity will have to be recalculated. Updates tend to occur more and more frequently, and so will the number of times the processor capacity has to be redimensioned.

When telecom services become more diversified, the difference in processor load between different types of call will increase. This means that even minor errors in estimates of the traffic mix will affect call handling capacity and thereby cause a risk of congestion, or inefficient use of resources.

To reduce the risk of inadequate accuracy in dimensioning processor capacity for AXE 10 exchanges, Ericsson has devel-

Fig. 1
The user interface to ACCIS - a Unix workstation with a TCP/IP connection to the central database. The users are currently spread among 11 Ericsson companies in as many countries

MAGNUS HÖGLUND
BJÖRN KIHLBLOM
Ericsson Telecom

oped a versatile and easy-to-use dimensioning tool called ACCIS, to be run on a workstation. The user interface is shown in Fig. 1.

Dimensioning for call capacity

The central processor normally sets the limit to the call capacity, Fig. 2. The number of regional processors can always be dimensioned to match any given call capacity figure. The unit of processor call capacity is Busy Hour Call Attempt, BHCA, which indicates the number of call attempts (successful or not) the processor can handle per hour. Other call capacity units are Busy Hour Call, BHC, and Busy Hour Call Completed, BHCC, which both include only successful calls. In calculations, calls per second is another common unit.

The basic issue when dimensioning an APZ processor is to determine the number of BHCA the processor can handle for a certain traffic profile. It is also essential to investigate the impact on capacity of variations in the traffic profile to cater for different scenarios; for example, extended use of call waiting and call forwarding services, or an increased use of a certain type of charging. To select the proper processor, the different types of APZ also have to be compared – with respect to their capacity for different traffic profiles. Not only new installations have to be dimensioned but also existing exchanges, for which a service upgrade is considered or a change in the traffic profile is expected.

Furthermore, ACCIS is used in product development to supply information about capacity requirements for new services and designs.

The traffic profile

The ACCIS user describes the traffic profile by specifying the mix of different types of call that is to be handled by the exchange. For a local exchange, the basic types are incoming, outgoing and internal calls, with specified signalling systems and types of subscriber. For each type of call, certain questions presented by ACCIS have to be answered; for example, the number of received and analysed digits and the intensity of each type of call in per cent of the total number of calls. There are also options to be specified, such as charging and various types of subscriber service. Moreover, the calculation requests a security margin – the 'traffic peak margin' – the level of which has to be specified.

ACCIS can handle all telephony system versions (APT) used in AXE 10 and all AXE 10 applications; for example, IN, GSM and ISDN. Traffic mixes are easily put together and varied. The data structure of the system permits virtually any traffic profile to be represented in a convenient fashion.

The output of the system is primarily the number of BHCA that a certain APZ processor can handle with the given traffic profile and traffic peak margin. Other information includes the permitted increase in traffic and the load for separate functions or parts of functions, giving the user information about e. g. the load increase per call caused by the use of Toll Ticketing instead of Pulse Meter charging, the load

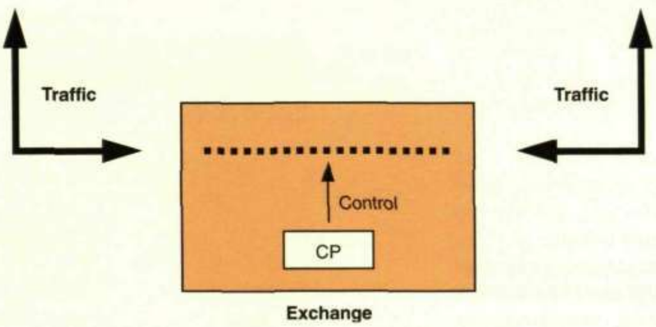


Fig. 2
The central processor is normally the primary object of call capacity dimensioning

for a credit card call compared with a normal call, or the effect of increasing the percentage of calls from, say, 20% to 100% using Calling Line Identification.

Basic theory of call capacity

Load components

For dimensioning purposes, the capacity of the central processor is divided into parts related to its different tasks.

To start with, 100% of the processor capacity cannot be taken into account. A limit to the processor's call handling capacity is set by requirements placed on real-time performance and behaviour in overload situations, Fig. 3. This limit is called processor loadability.

Box A

Example of a traffic mix

Creating a traffic mix in ACCIS is a straightforward matter. For each application, there is a list of functions, types of call, etc, from which the wanted functions are selected. Each function in the mix is then specified by answering questions put by the system. These questions correspond to the parameters of the formula for the respective functions. Figures for the traffic peak margin and usage load are also prompted. Fig. A illustrates a simple example of a traffic mix for a local exchange.

The information given by the user is underlined. The result will indicate that, for a specific processor, this traffic mix gives 360,000 BHCA if the traffic peak margin is applied, and 451,000 BHCA if not.

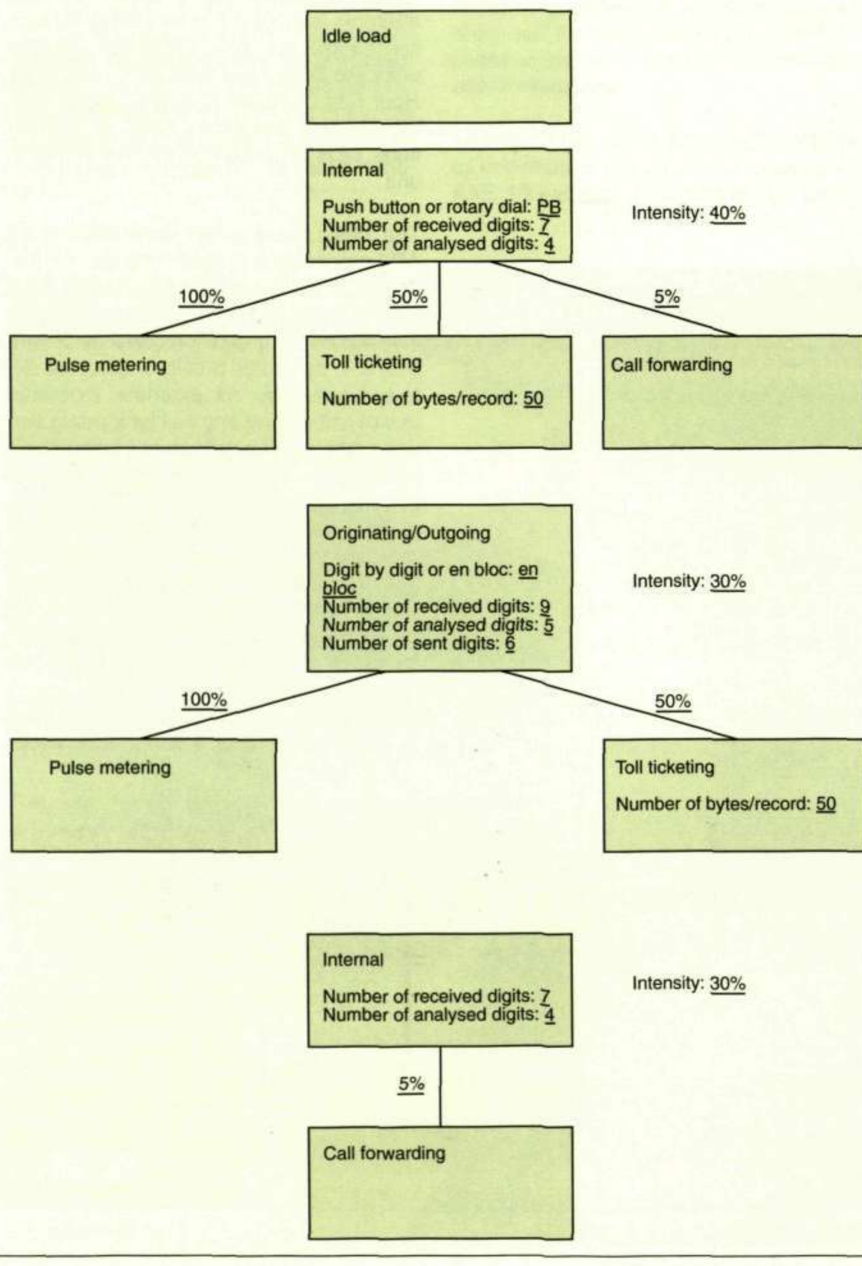


Fig. A
An example of a small mix containing only three types of call. The input data, given by the user, is underlined

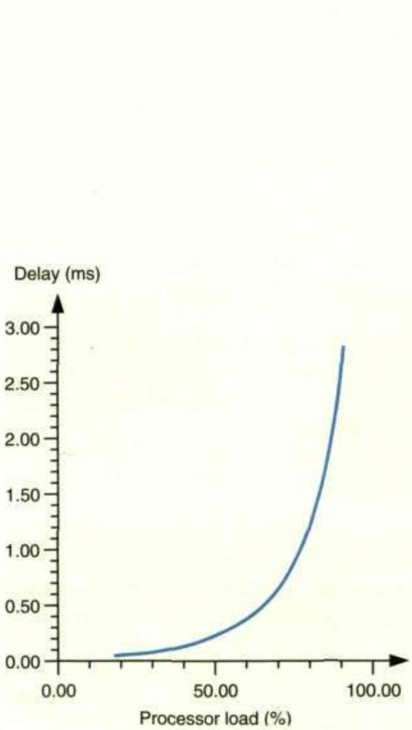


Fig. 3
The loadability limit is set by the requirements for limited real-time delays in the system. The graph shows how the internal processor delays at a specific execution level will increase towards infinity as the processor load tends towards 100

The processor loadability varies for the different versions of the APZ processors. It is usually between 90 and 97%; the latter figure applying to the more powerful processors. In a few cases – for example, in exchanges in which a major part of the traffic is operator-handled – loadability will also be affected by the type of traffic.

The first processor-load component to be considered is called the idle load. This is the load on the processor when there is no traffic through the exchange or no other external activity going on. This load is caused by jobs initiated by job-table signals, for supervision of software and hardware resources. The idle load varies with the type of APT and APZ and with the size of the exchange.

The second component is a reservation of processor time for execution of commands; it is referred to as usage load. The usage load, expressed as a percentage of the processor's total capacity, is usually in the range 2 - 6%, depending on processor type and type of exchange.

Subtracting idle load and usage load from the loadability gives the processor capacity available for traffic handling, Fig. 4. Typical figures for small processors are a loadability of 95%, an idle load of 17% and a usage load of 3%, leaving 75% of the processor's total capacity for traffic handling.

Traffic handling

Traffic handling includes all call-related program execution by the processor, such as call set-up and disconnection, subscriber and network services, charging, etc.

In dimensioning, focus is on peak-load situations. The hour of the day with the highest average load is defined as the 'busy hour'.

In the example above, the load figures are given as a percentage of the total capacity. A more convenient way is to use milliseconds of execution time per second, converting 1% of the capacity of the processor to 10 ms/s. Thus, in the example, 750 ms/s are available to the calls that constitute the traffic.

However, it is necessary to reserve processor time for variations within the busy-hour traffic. This spare capacity is referred to as 'traffic peak margin' and expressed in per cent of the execution time available for traffic handling, Fig. 5. If the traffic peak margin in the example is set to 35%, then 750/1.35=556 ms/s are available to the estimated busy hour traffic.

Another means of expressing this safety margin is the dimensioning factor defined as 1 + the traffic peak margin. Thus, a traffic peak margin of 35% is equivalent to a dimensioning factor of 1.35. CCITT, Q.543, recommends a dimensioning factor of 1.2 and 1.35 for transit and local traffic, respectively.

For a simplified example concerning a local exchange, the following traffic mix is assumed:

Type of call	p_i	t_i
– Originating/ outgoing	35%	16.13 ms/call
– Incoming/ terminating	35%	11.27 ms/call
– Internal	30%	12.87 ms/call

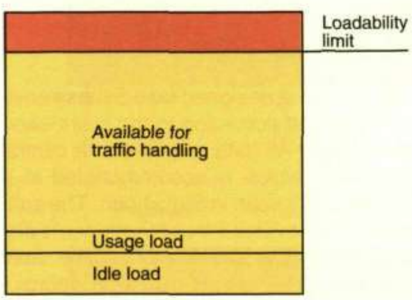
Where:
p_i = portion of calls that are type i calls.
t_i = execution time for a type i call.

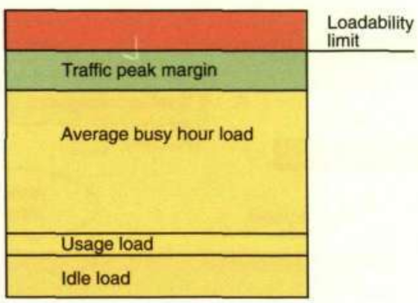
Loadability	90%
Idle Load	22%
Usage Load	6%

This gives an average execution time $T=\Sigma p_i t_i$ of 13.45 ms/call and 620 ms/s execution time available for traffic handling. Leaving out the traffic peak margin, the traffic handling capacity is 620/13.45 = 46.1 calls/s (16 outgoing, 16 incoming, and 14 internal calls per second). To obtain the

Fig. 4, below left
The constituents of the processor capacity. The loadability limit is set by real-time requirements. The idle load is the load on the processor without any enforced tasks, and the usage load is the capacity reserved for operation and maintenance

Fig. 5, below right
The traffic peak margin takes into account the traffic variations within the busy hour and is expressed as a percentage of the average busy hour traffic





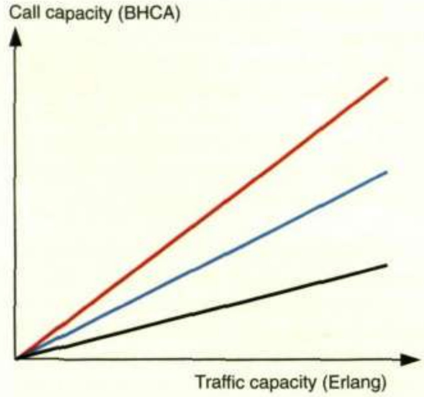


Fig. 6
The relation between traffic and call capacity for different holding times. The diagram shows the formula BHCA = 3600 x holding time x Erlang, for three holding times

Holding time 60 s
Holding time 90 s
Holding time 120 s

capacity in BHCA – Busy Hour Call Attempts – the figure 46.1 calls/s is multiplied by 3600, giving 166,000 call attempts per hour. The term BHCA indicates that unsuccessful calls and service requests are also included in the figure.

The capacity will be reduced through the introduction of a traffic peak margin. With a margin of 35%, the time available for normal busy hour traffic is 556 ms/s. This gives a capacity of 459/13.45 = 34.1 calls/s, corresponding to 123,000 BHCA.

Processor capacity versus call capacity

The figures used in the example are valid for a control system for small exchanges. If a more powerful processor is assumed, 3.5 times faster per instruction compared with the previous one, then the table in the previous example will be as follows:

Type of call	p_i	t_i
– Originating/ outgoing	35%	4.61 ms/call
– Incoming/ terminating	35%	3.22 ms/call
– Internal	30%	3.68 ms/call

Also, the idle load and the usage load will be changed although by different factors, here assumed to be 3.2 and 2,0 respectively. The idle load is then reduced from 22 to 6.9% and the usage load is reduced from 6.0 to 3.0%. Furthermore, due to shorter average jobs, the loadability can be increased to 95%.

The processor time available for traffic handling will then be 95.0–3.0–6.9 = 85.1% = 851 ms/s, and the load for the average call 3.84 ms. Thus, the number of calls per second, excluding the traffic peak margin, is 851/3.84 = 221 calls/s, corresponding to 797,000 BHCA. Including a 35% margin, the result is 590,000 BHCA.

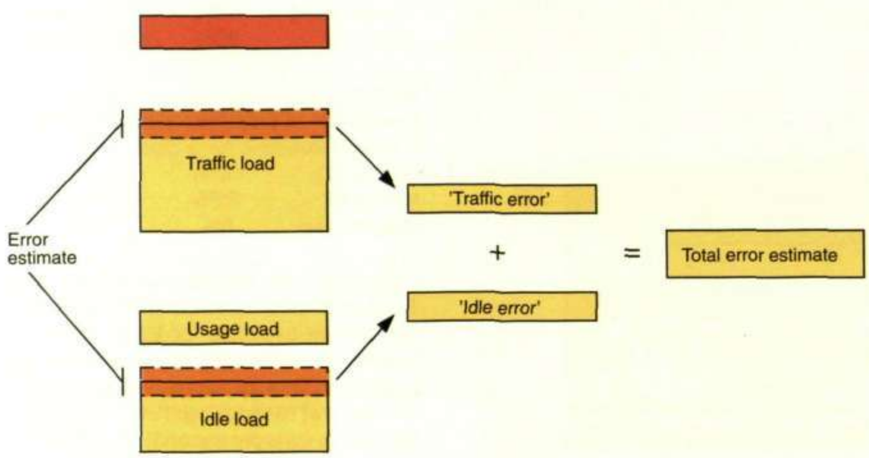
The ratio between this capacity and the capacity of the smaller processor is 4.8, which is more than the 3.5 times reduction in instruction time. The reason is that, with a more powerful processor, both idle load and usage load are reduced, which means that the ratio of time available for traffic handling is increased.

ACCIS also accepts traffic figures given in Erlang, holding time per call and number of calls received per second, Fig. 6.

Error estimates

The load figures for the individual call-types are based on estimates or measurements of varying accuracy. In ACCIS, an error estimate is assigned to each load figure for calls and the presented result includes an estimated error margin, Fig. 7.

Fig 7
An automatic error estimate is part of the dimensioning result. The total error estimate is based on a weighted average of the estimated error for each function constituting the traffic mix



System structure

The system is designed for a Sybase environment and according to the client-server concept. All data is stored in a central database, which is accommodated in a server at Ericsson in Stockholm. The software that provides the user interface is distributed to the workstations. The user needs a UNIX workstation with approxi-

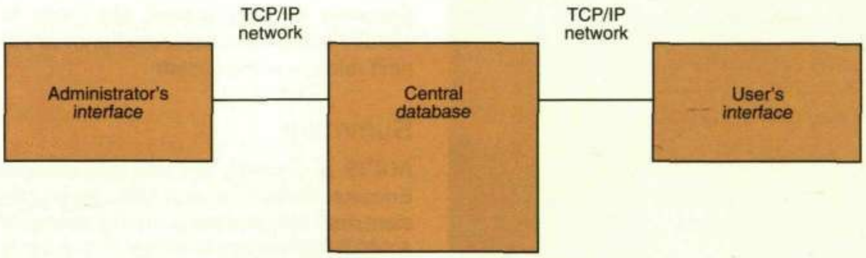


Fig. 8
The central database is accessed via two different interfaces. The user's interface contains the necessary facilities for entering traffic mixes and making capacity calculations. The administrator's interface provides the facilities for entering and updating basic data for the ACCIS functions

mately 3.5 Mbyte of memory available for the ACCIS software. The TCP/IP protocol is used for communication, and a connection to the Ericsson TCP/IP network is therefore needed.

There are two interfaces: one for the system administrator and one for the user, Fig. 8. The administrator enters the basic information for each type of call, such as load formula and questions that are to be presented to the user. He also administrates the system and can allow access to certain parts of the database for certain users; for example, traffic profiles specified by a user will be available to this user only. The structure of the database also allows the administrator to choose alternative presentations adapted to the user's need. The user interface contains facilities for producing and manipulating traffic profiles.

Sources of call capacity dimensioning

There are mainly three types of data needed for call capacity dimensioning:
- Data per type of APZ
  - Loadability limit
  - Execution time per instruction
- Data per Application System
  - Formulas for the execution time per call
  - Idle load formulas
- Data per Exchange
  - Expected traffic mix
  - Expected usage of functions
  - Exchange data, e.g. the number of lines

The third type, data per exchange, is frequently changed in exchanges in operation.

Processor load data

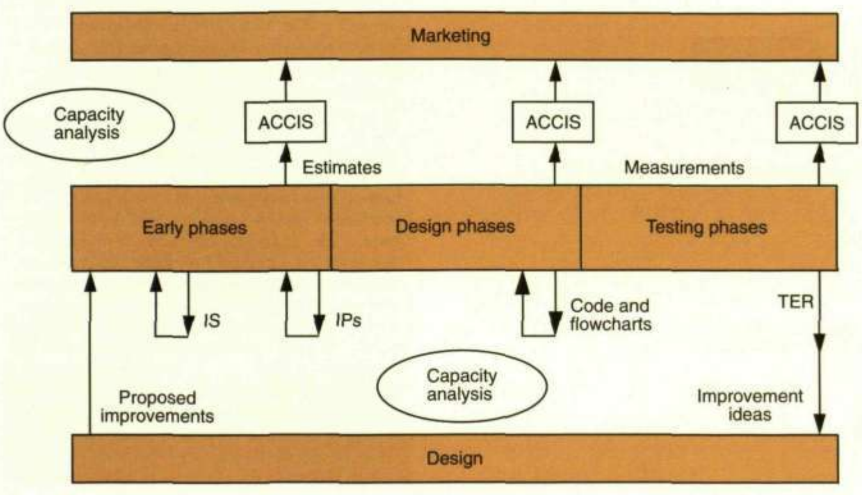
Processor load data is collected mainly for two purposes: Firstly, to create a base for dimensioning the control system in an AXE for new or upgraded applications or for changes in the traffic situation, Fig. 9. Secondly, to identify applications that use excessive processing power, in order to continuously improve the AXE 10 call handling capacity.

Data for dimensioning
During the early design phases for new functions, the capacity impact is described, and the overall effects are brought into ACCIS for preliminary capacity dimensioning. The first measurements of call capacity are made in a simulated environment during the early testing phases. Supplementary measurements are made at an AXE test plant, Fig. 10. The results are then used in ACCIS.

Data for design
Processor load data is fed back to the designers and the project management, in order to supervise the processing power required by the applications designed. Short-term feedback, for example in the form of quality inspections, aims at identifying and reducing excessive use of processing power during the design work.

Fig. 9
Development project activities of controlling and supervising the call capacity of a new system. The vertical arrows represent the points in time during design or testing, when processor load data is extracted. The data is analysed and transferred to the marketing or design office, or fed back into the development project

IS	Implementation sketch
IP	Implementation proposal
TER	Test report



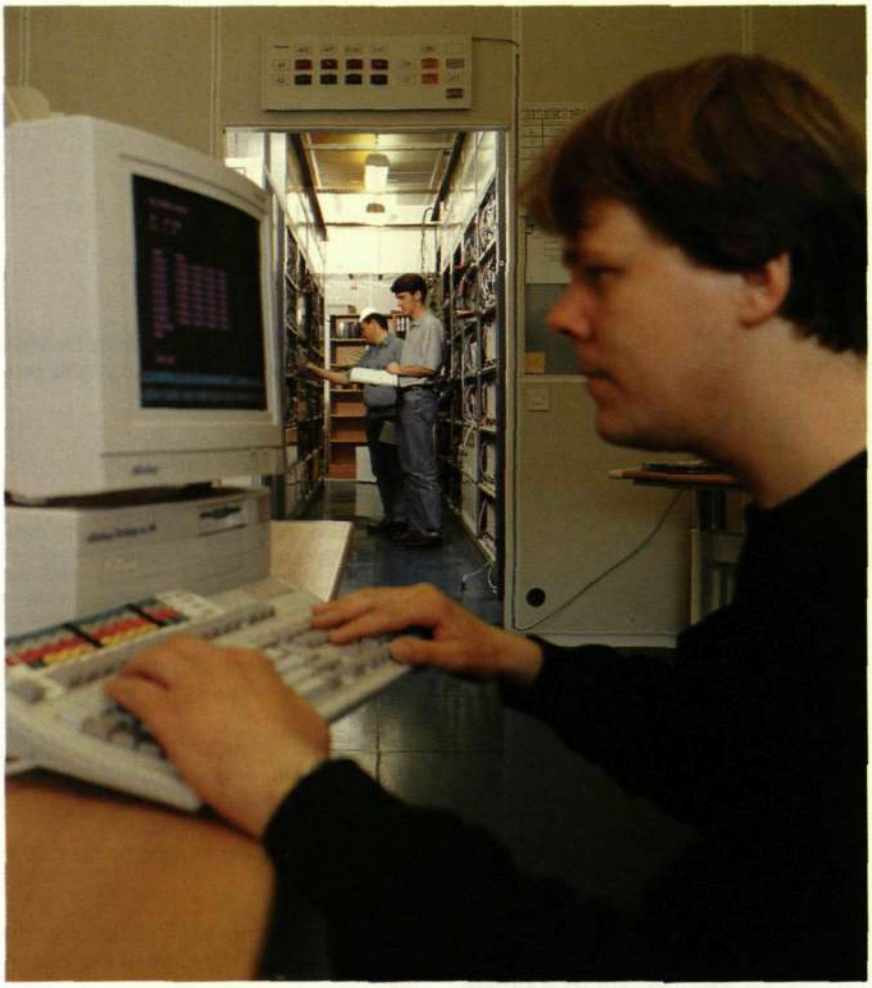


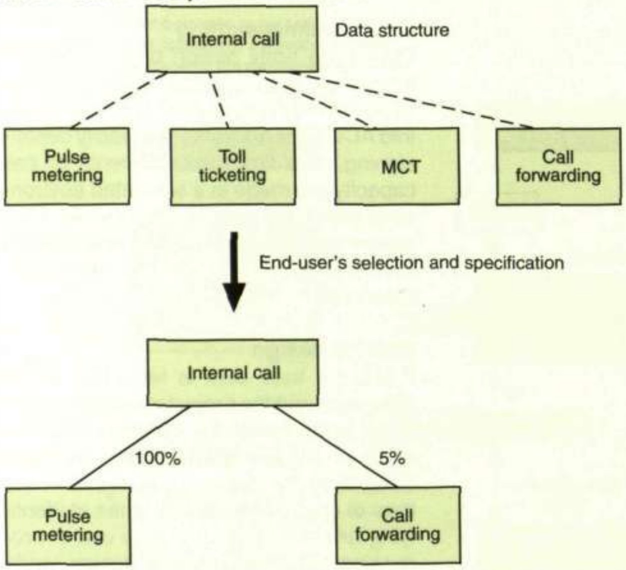
Fig. 10
The execution time for various types of calls are measured at an AXE test plant using internal measurement functions. An external program adapts the result to ACCISS format and also supplies detailed data for design feedback

Separate capacity studies are made to identify areas to be corrected prior to the next release of the system.

Summary

ACCIS is currently put into operation at Ericsson companies worldwide as an efficient tool for improved dimensioning of AXE 10 exchanges. The risk of oversights and mistakes in the dimensioning work is reduced, since ACCIS prompts the user to enter the required input data. The availability of a database for dimensioning data also makes it easy to check if the processing power in an existing exchange is sufficient when traffic grows or new features are introduced. With ACCIS, the processing power can be conveniently dimensioned with accurate results, based on the most recent data available.

Fig. B1
The administrator sets up possible options and the user selects and specifies his functions



Box B

Data structure

The data structure of ACCIS is designed to cater for any foreseeable type of traffic profile. Information is presented in a form which is convenient to the user. The basic unit of the structure is the 'function', which is an item that stores load information entered into the system. A function represents a certain task performed by the application system. The task may be part of a traffic profile, e.g. the set-up of an internal call, call forwarding, or charging. The function, defined by the administrator, contains the load formula and other information necessary for the functionality it represents. The load formula is used to calculate the execution time in ms when the associated task is executed in a given processor.

An example of a formula for an ordinary internal call could be: execution time = $17+0.3p_1+0.5p_2$, where the parameters p_1 and p_2 correspond to the number of received and analysed digits respectively. (The figures do not relate to any specific version of APT or APZ.) A number of factors related to the different versions of APZ are assigned to each function. Since the instructions to be performed by the processor are different for different functions, the same factor cannot be used for all functions.

Furthermore, it is not convenient to describe traffic profiles with one function for each type of call. Therefore, the system allows for a function to have

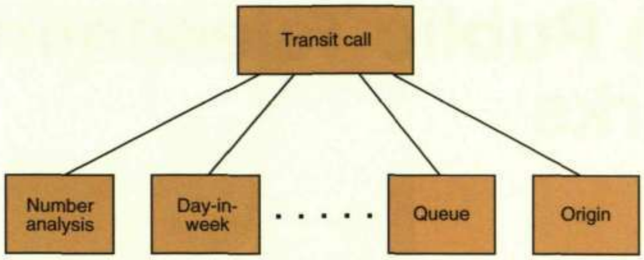


Figure. B2
The administrator provides a set of service modules from which the user selects those he needs for his IN service

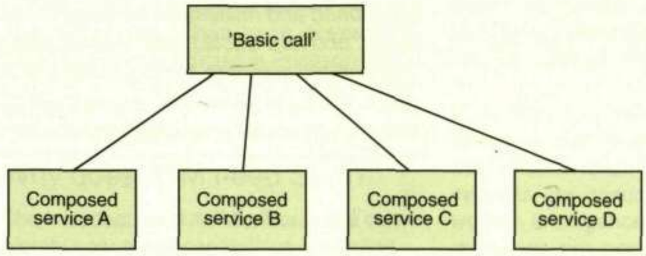


Figure B3
The administrator can provide the user with a set of complete IN services, which means that the user need not concern himself with the individual service modules, of which the service consists

other functions attached to it, as illustrated in Fig. B1. For instance, the function internal call may have subordinated functions specifying charging and subscriber services, for example. Such functions have their own load formula and parameters. From the set of functions that can be attached to a certain function, as given by the administrator, the user selects those that are relevant and states the frequency in per cent for the use of each of them, Fig. B1.

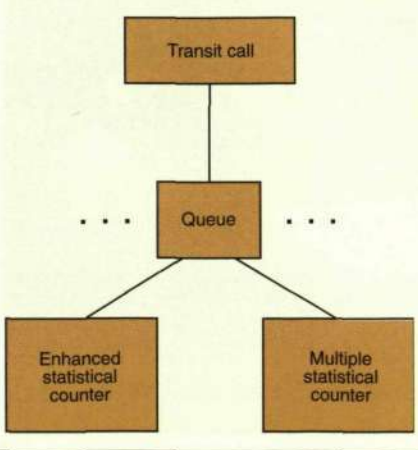
An intelligent network (IN) application may serve as an illustration of the flexibility of the system. IN services are obtained by combining a number of service modules, each of which performs a certain task. The modules can be combined in different ways to perform a wide range of services. The ACCIS administrator can set up a number of functions - each representing a service module - out of which the user can make combinations. In an SSCP for instance, the IN services are often executed together with an ordinary call, such as a transit call. Fig. B2 shows a possible solution for the setting-up of functions for a transit call and IN modules.

However, representing a service as a combination of individual service modules may be too detailed in the initial projecting phase, when the exact set of modules to be employed has not yet been decided and only a rough estimate is asked for. In this case it may be more appropriate to provide a condensed set of functions, each representing a complete IN service, Fig. B3.

Two levels of functions will normally suffice, but the system allows for an expansion of the number of levels. A third level may be convenient, for example when a certain function can occur together with certain other functions only. In the example given in Fig. B4, the service modules 'Enhanced statistical counter' and 'Multiple statistical counter' can only be used together with the 'Queue' module.

With the structure shown in the figure, virtually any traffic profile can be represented. The structure also allows the administrator to present the different types of call with a suitable level of details for the convenience of the user, as more accurate information is obtained from the design and testing processes. The load formula and other information can be refined through an up-dating routine included in the system. The user responds to the changes in the formula, but may otherwise keep his traffic mix unchanged.

Figure B4
For some IN applications, it may be convenient to arrange the service modules in two levels. When the module 'Queue' is selected, the user has the option to attach 'Enhanced Statistical Counter' and 'Multiple Statistical Counter'



ATM in Public Telecommunications Networks

Bengt Lagerstedt, Hans Nyman

SDH and ATM will both be crucial to the successful implementation of commercial broadband and multimedia services in public telecom networks. But how do they interrelate, and how should they best be deployed to maximise business opportunities? The authors compare the capabilities of SDH and ATM systems, examine how changes in the commercial environment will affect the introduction of systems and services, and present possible scenarios for future network evolution.

The next few years will see major changes in the way transport networks are viewed, built and operated.

Many network operators are now about to deploy SDH (Synchronous Digital Hierarchy) systems throughout their transport networks. At the same time, there is talk of introducing ATM (Asynchronous Transfer Mode) – either on top of SDH, or possibly even instead of SDH, using existing PDH (Plesiochronous Digital Hierarchy) transmission resources.

How do these technologies – SDH and ATM – relate, and how should they be introduced into telecom networks?

ATM — the 'ultimate' transmission technology?

One major goal for public telecom network operators is to find a single technology to handle the transport and switching of all types of service within a common infrastructure.

ATM has the potential to provide this solution, creating a bearer network with unlimited bandwidth 'granularity' (i.e. the bandwidth of a call can vary freely, even throughout the duration of the call), and capable of coping with very high bandwidth connections.

ATM standards are now sufficiently well-defined to allow a useful degree of interworking between systems. Component design has reached the point at which ATM switching products can be manufactured economically, and the demand for the services that ATM will support already exists.

Ericsson has recently introduced its own ATM product offering: a completely scalable and service-independent platform suitable for use in any application – from a

Fig.1
Computer-based remote teaching is possible with the aid of broadband technology in the telecommunications network. Broadband with ATM technology is opening up many new application areas for telecom networks

BENGT LAGERSTEDT
HANS NYMAN
Ericsson Telecom

floor hub in a customer premises network to a large transit node in a public network. The system itself has been covered in detail in previous articles in Ericsson Review[1, 2] and is briefly described later in this article.

Why does ATM need SDH?

The introduction of ATM will create a completely new set of capabilities, including
- common switching and transport for all services
- increased bandwidth granularity, with bandwidth allocation from virtually zero to very high bit rates
- support of variable-bit-rate services
- support of multimedia services.

It is because of these features that ATM has been chosen as the core technology for broadband ISDN.

However, ATM does have some disadvantages. There are transit delays for low-speed isochronous services, which have led some to question its ability to handle voice-type services cost-effectively, without the need for echo cancellers. ATM also adds complexity to the network and introduces new performance parameters, such as cell-loss and congestion, which operators will have to learn how to handle.

Both SDH and ATM represent a convergence of the 'traditional' switching and transmission roles. Traditionally, switching systems are assumed to contain a network's intelligence, while transmission resources are passive conduits through which switched signals flow.

Existing transmission networks are optimised for voice-type traffic: low-bandwidth services that, when aggregated, follow fairly static and predictable traffic patterns. Present-day PDH transmission systems were introduced to reduce costs by providing higher capacity while utilising existing cable resources. However, this has led to a very rigid network, costly to maintain and with uncertain performance.

An ATM-based network can be built using either PDH or SDH – or both – as the physical media. Another alternative which has also been proposed is to use pure ATM as the bearer. However, standards have yet

BOX A

ATM cell structure

An ATM cell has a fixed length of 53 bytes, or octets, divided into a 5-octet header and a 48-octet information field (payload), Fig. A.

As indicated in Fig. A, the ATM cell header is structured as a number field. The main function of the header is to route the cell from the point of origin to the point of destination. The routing information is contained in the VPI (Virtual Path Identifier) and VCI (Virtual Channel Identifier) fields. As shown in the figure, the header is slightly different in the UNI (User-Network Interface) compared with the NNI (Network-Node Interface). The UNI contains 4 bits for GFC (Generic Flow Control). It will be used to ensure fair and efficient use of available capacity between terminal and network.

The Payload Type field, PT, is used to indicate whether the cell contains user information or special network information, e.g. for maintenance.

The Cell Loss Priority field, CLP, indicates a two-level priority and is used if it becomes necessary to discard cells, depending on network conditions.

The header information is protected by a check sum, which is contained in the HEC (Header Error Control) field.

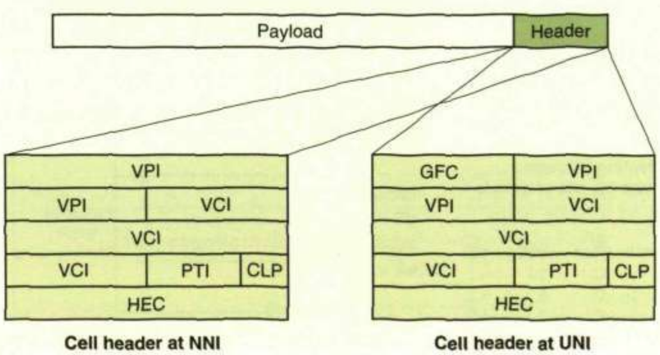


Fig. A
The length of the ATM cell is 53 octets. The cell header occupies 5 octets and is structured as a number field. The cell header contains routing and traffic control information. The structure of the cell header is slightly different at the user network interface, UNI, compared with the network node interface, NNI

GFC	Generic Flow Control
VPI	Virtual Path Identifier
VCI	Virtual Channel Identifier
HEC	Header Error Control
PTI	Payload Type Identifier
CLP	Cell Loss Priority

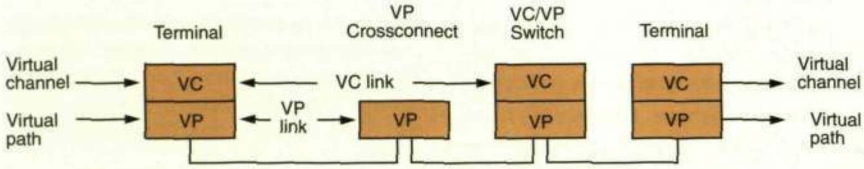


Fig. 2
A VP/VC path consists of a number of interconnected VP/VC links. Switching/cross-connection can be performed either at the VP or VC level. The VPI and VCI defines a two-tier handling and routing structure

to be developed for pure ATM transmission, and the technique would place severe limitations on the use of the network in order to avoid multiple conversions between ATM and STM (Synchronous Transfer Mode), and the consequent degradation of performance.

As a bearer for ATM, only SDH can offer the flexibility and performance monitoring needed to handle both ATM and existing traffic efficiently. PDH is less flexible and besides, PDH facilities are often fully utilised by existing traffic

ATM on top of SDH represents perhaps the ultimate convergence of switching and transport – providing both a smooth migration path and the service-independence needed to support future services, such as multimedia communications.

Basic characteristics of ATM

Four essential characteristics should be highlighted here, because of their importance to future public telecom networks:
- bandwidth and media flexibility
- multiplexing and cross-connect/switching capabilities
- the 'core-and-edge' principle
- management capabilities.

Bandwidth and media flexibility
The basic information transfer unit in ATM is a small, fixed-size packet, known as the cell (see Box A). This permits the information transfer rate to adapt to the actual service requirement. Depending on the capacity required, the number of cells per time unit can be increased up to the transmission bit-rate limit set by the physical medium.

In addition to data cells there are cells for signalling and maintenance and idle cells. Signalling cells are used between an end user and the network, or between nodes in the network. Their function is to set up a service, e.g. a connection. Maintenance cells provide supervision of the ATM layer (the F4-F5 flows described later). The idle cells are used to fill the transmission capacity up to the rate of the physical medium.

The flexibility of ATM enables it to use existing infrastructure for physical transport, as mentioned above. The preferred transmission medium is SDH using VC4 (155 Mbit/s), but lower capacity containers are also defined. The PDH transmission hierarchy can also be used (2, 34, 140 Mbit/s transmission systems), provided the performance of these systems is adequate.

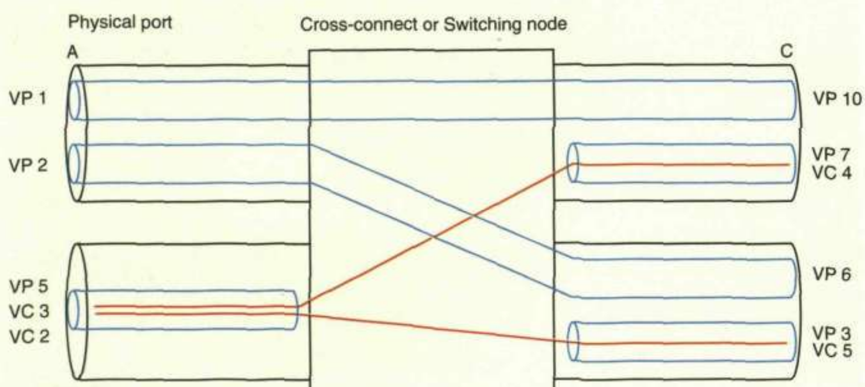


Fig. 3
***VP and VC cross-connection and switching.* The VPI and VCI values are valid for each link only. In each cross-connect/switch, new VPI/VCI values are assigned to the cell. The combination of physical port and VPI/VCI values provides identification of the cell. Routing through a switch is then performed with the aid of translating tables**

Port	VPI	VCI	connected to	Port	VPI	VCI
A	1	–		C	10	–
A	2	–		D	6	–
B	5	3		C	7	4
B	5	2		D	3	5

Multiplexing and cross-connect/-switching capabilities
The ATM cell is the basic multiplexing unit. Each cell (or information unit) contains its own connection and routing information. This enables direct multiplexing or demultiplexing of service channels, where each channel may carry different bit rates.

ATM cells are identified and routed by information contained in the header. This routing information is located in the VPI (Virtual Path Identifier) and VCI (Virtual Channel Identifier) fields. The VPI and VCI values define a two-tier handling and routing structure, Figs. 2 and 3.

A Virtual Path (VP) is a bundle of multiplexed circuits between two termination points (e.g. switching systems, LAN gateways or private network gateways). The VP provides a direct logical link between virtual path terminators. The VPI value identifies the VP.

The VP concept allows multiple Virtual channels (VC) to be handled as a single unit. VCs with common properties (e.g. quality of service) can be grouped together into bundles that can be transported, processed and managed as one unit. The flexible bundling simplifies operation and maintenance.

Both VPs and VCs can be used to provide semi-permanent paths in the ATM network. Routes are established or released, from an operations support system, by the setting of 'path connect tables' in the cross-connect equipment or multiplexers on the path. Required transmission capacity is also reserved in the same way.

BOX B

Maintenance capabilities in an ATM network

To maintain the level of quality in the network, CCITT has defined a hierarchy of maintenance flows, defined for each transmission level as indicated in Fig. B. These maintenance flows provide the capabilities to quickly detect degradation in the network, detect a faulty element, isolate it and provide information to maintenance staff.

Five levels have been defined. The lower three concern the physical layer, while the upper two concern the ATM layer.

Physical layer
The first level, F1, applies to the regenerator section and is mainly used for fault localisation of regenerator sections. The second, F2, applies to the digital transmission section. It is used to monitor the section and to provide protection switching. The third level, F3, applies to the digital path and provides supervision between ATM nodes.

The F1 to F3 flows are used by the network operator. For SDH, the flows correspond to regenerator, section and path overhead and are standardised by CCITT. For PDH, the availability of F1 and F2 depends on the equipment used and is specific for each brand of system. The F3 flow is carried in bits added to the PDH frame.

ATM layer
The fourth level, F4, is used to monitor the VP between the user terminal and the switch and between switches. Supervision is achieved by inserting recurrent ATM OAM (Operation and Maintenance) cells.

The fifth level, F5, is used to monitor the VC between two user terminals. In this case too, supervision is achieved through OAM cells. The F5 flow is also used to indicate failures in the AAL (ATM Adaptation Layer) between terminals.

The F4 and F5 flows are oriented towards control of the quality of service and can be used by both the operator and the end user.

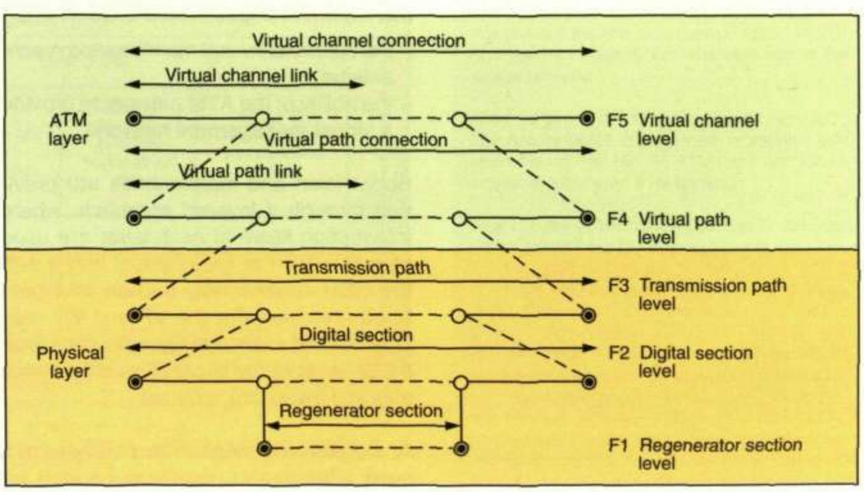
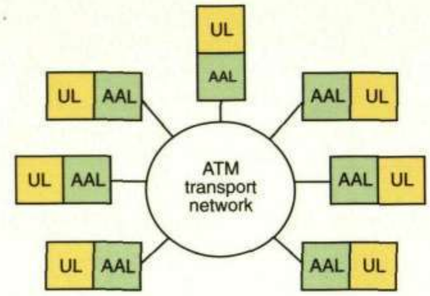


Fig. B
Maintenance flows at five levels are defined. The lower three relate to the physical media and depend on the capabilities of the PDH or SDH transport system. The upper two relate to the ATM level and depend on ATM maintenance fields

Fig. 4
Core-and-edge principle. Adaptation between the user layer, UL, and the ATM transport network is done through the service dependent ATM adaptation layer, ALL. The transport of the cells will thus be service-independent, and different services can dynamically share the transport network



VCs can also be used for on-demand switching. Connections are established by signalling, between user and network and within the network, using ATM cells.

The 'core-and-edge' principle

An important characteristic of ATM is related to its protocol architecture, built around the so-called 'core-and-edge' principle.

The protocol functions specific to the information type transported (such as retransmissions, flow control, delay equalisation) are performed in terminals at the edges of the ATM network, Fig. 4.

This leaves an efficient, service-independent core network, including only simple cell-transport and switching functions. In the ATM nodes in this core, there are no error checks on the information field; nor are there any flow controls. The cell information is read, the HEC is used to correct single errors that might affect the address, and the cell is then switched.

The ATM Adaptation Layer, AAL, is used at the edge of the network to enhance the service provided. The AAL, which includes service-dependent functions, is described in Box C.

The 'core-and-edge' principle makes it simple to introduce new services. The specific service-dependent functions are handled at the network terminal: the network itself only has to handle cells.

Management capabilities

Two important characteristics of ATM management capabilities are:

- the supervision and maintenance mechanisms
- the ability of the ATM network to provide a virtual management network.

Supervision and maintenance are provided through a layered approach, where information flows at each layer are used (see Box B). For the physical levels 1-3, the SDH regenerator, section and path flows are used. For the VP and VC sublayers, ATM maintenance cells are used. If PDH is used, the capabilities of the lower physical levels are reduced.

At the planning stage, transport networks must allocate the resources needed for their management. The requirements often being uncertain, a substantial margin has to be allowed for. In the ATM logical network there is the advantage of flexibility. Since management information can be transported in the cells within the network itself, only minimal resources need to be allocated. As the requirements on security and capacity grow, new VP/VCs can be added and the assigned capacity increased.

The changing commercial environment

Telecommunications continues to be a fast-changing and growing market. Deregulation and liberalisation are creating competitive marketplaces, and causing changes in the behaviour of both users and providers.

Users are demanding better quality of service, broader bandwidths and more features at lower cost. They are learning to shop around for the best deal.

New, alternative network operators are emerging. Utilities like railway companies and electricity boards are building fibreoptic networks that can be used to provide new or competing services to end users.

To be competitive in this environment, existing network operators are beginning to reorganise. This reorganisation has two goals: meeting the demand for advanced services from the large and lucrative business customers, and creating an 'internal market' between units in the network operator's organisation, which will tend to maximise efficiency and minimise costs.

The end result of this reorganisation may be that network operators atomise into business units. A 'transport company' would concentrate on the core network service of bandwidth provision. Service providers (both within and potentially outside the network operator) would 'add value' by buying bandwidth from the 'transport company' and reselling it, having added services, to end users, Fig. 5.

The service providers must be able to react quickly to changes in demand. New ser-

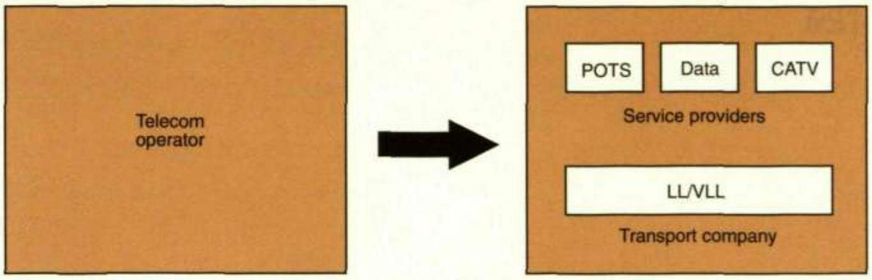


Fig. 5
Telecom operators are starting to operate in a way that focus on the needs of specific customers, services and the basic network

vices must be implemented quickly – and discontinued quickly if they prove to be unprofitable. The emphasis is on short-term investment, with short-term payback.

In contrast, the 'transport company' has a long-term interest in the network infrastructure. All services need bandwidth. It can therefore make long-term investments, to create an infrastructure with the necessary flexibility and future-proofing to support on-demand bandwidth provision with high granularity.

ATM will play a major role in both these activities. In the near term, it can be used as a vehicle for adding value to the existing transport network. In the longer term, ATM can be used to build a ubiquitous network, to handle transport of all types of service.

Deployment of ATM in the network

The first use of ATM technology will be within the Customer Premises Network, CPN. Here, it will first support high-speed data communication in and between LANs, and later be used as an infrastructure common to all services in the CPN, including voice and video communication as well as data. The first multimedia applications will therefore arise in private networks.

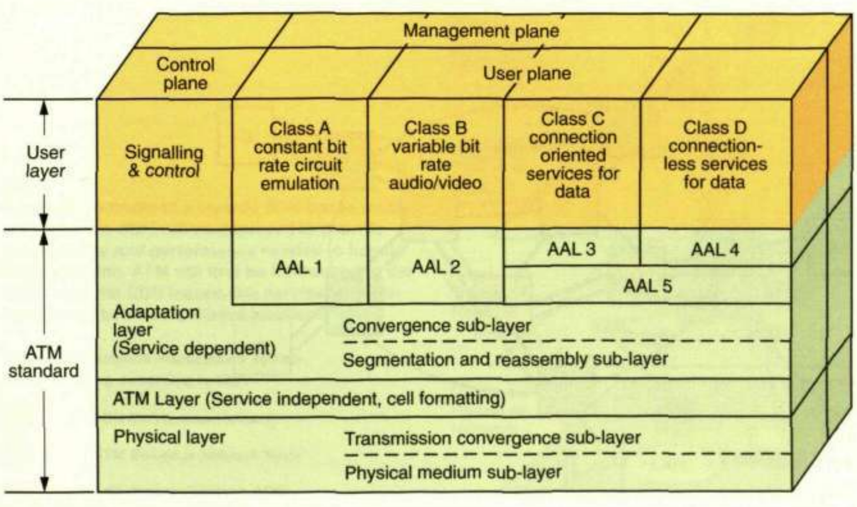
In this scenario there is no immediate need for ATM capabilities in the public network; instead, pipelines to carry ATM traffic between sites can be set up using SDH – or PDH – transport network systems.

However, network operators can add value by introducing their own ATM nodes into the public network. The first service to be provided will probably be the Virtual Leased Line (VLL) service.

Current leased-line services are inflexible: only constant bit rates at hierarchical levels can be provided. A VLL service based on the VP concept offers the following improvements:

Tailored VP capacity
The capacity can be directly tailored to customers' needs and can easily be changed without modifying the interface structure.

Fig. C
The B-ISDN reference model shows the variety of service classes supported and the layers of the ATM standard



BOX C

ATM layers

There are three layers in the ATM standard as shown in Fig. C. The first layer is the physical layer defining the physical interfaces and framing protocols. Several options are available. The physical layer can be based on SDH or PDH transmission systems, for example.

The second ATM layer is independent of the physical medium. It defines the cell structure, provides multiplexing and demultiplexing and VPI/VCI translations. It defines how the cells flow in the logical network.

The third layer is the ATM Adaptation Layer, AAL. The AAL provides the important adaptation between the service and the ATM layer and allows service-independent ATM transport.

Four classes of services, A,B,C and D, are supported by the AAL. There are 5 types of AAL. AAL 1 – 4 support classes A – D respectively. AAL 5 is a more effective version of AALs 3 and 4 and supports classes C and D for high-speed data.

The AAL performs mapping between the original service format and the information field of the ATM cell. Some examples of the functions provided by the AAL are variable-length packet delineation, sequence numbering, clock recovery and performance monitoring.

The number of logical connections that can be offered to a user through the UNI (User Network Interface) is large (maximum 256 VPs and 67,536 VCs).

Tailored Quality Of Service, QOS
A tailored QOS can be offered, matching the service of the user. Multiple QOS classes and performance parameters can be selected. Voice services, for instance, require low transmission delays but can tolerate bit errors. Data communication, on the other hand, is tolerant of network delays but is sensitive to bit errors. The QOS level can be contractually agreed with the customer, and audited.

Reduced transmission cost
The sharing of network resources improves utilisation and reduces costs – these savings can be passed on to customers. It will also be possible to charge customers according to their use of the network.

Enhanced Operations And Maintenance (OAM) capabilities
OAM cells can be used to implement VP maintenance and provide performance monitoring, testing, path tracing, etc.

In the initial phase, bandwidth will be allocated to VLLs according to 'peak rate', with customer adherence to the agreed rate policed at the network terminal; and with charging according to peak rate. Customer management systems will permit the capacity to be increased or decreased on demand, using signalling over the ATM logical network.

As standards evolve, it will be possible to upgrade and offer bandwidth allocation according to 'mean rate', using statistical multiplexing in the network. Ericsson is actively involved in this standardisation work. Applying the statistical multiplexing method means that charging will more accurately reflect the use of network capacity.

With a VP-based private network, the user has the VCs available for internal routing and switching. However, the public network will also be equipped to route on VCs, thus providing enhanced service. Fig. 6 shows an example of a VC-based VLL network.

Other services, such as SMDS/CBDS and Frame Relay, can easily be added according to demand, by connecting servers to ATM nodes in the network.[1]

As time passes, demand will grow and equipment costs will fall. ATM will shift from being an 'overlay' mechanism for special services, to being a basic switching and transport mechanism throughout the network.

In the residential area, ATM technology could be used in the drive to provide new and enhanced entertainment services, such as video on demand, to the end user. The flexibility of ATM makes it possible to support a multitude of services, e. g. distance education, home shopping and games. The use of ATM for residential services will increase demand and speed up the transition from ATM applied in an over-

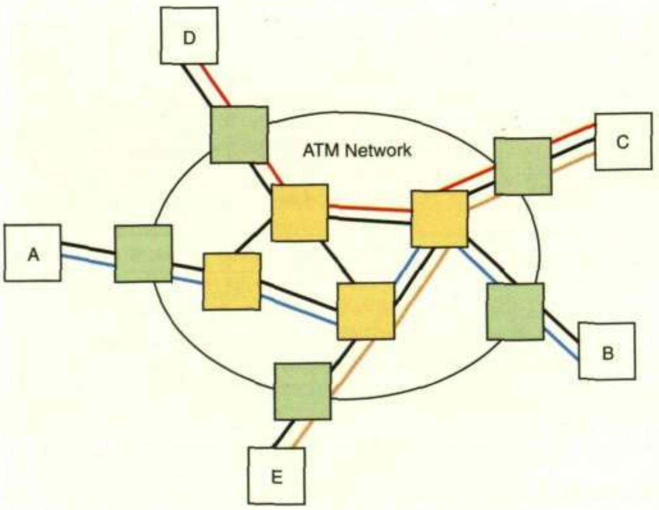


Fig.6
ATM network providing VLL service. The ATM network consists of a number of ATM cross-connects that can provide routing both at the VP and VC level. A flow enforcement function is located at the edge of the ATM network to protect the network against potential overload. This function ensures that no connection violates the conditions agreed when the connections were set up. ***Additional services can be implemented by*** **adding services to one or more of the cross-connect nodes**

ATM cross-connect node

Flow Enforcement (Police function)

Virtual Path
Virtual connection 1
Virtual connection 2

lay network for business services to ATM as a basic technology in the infrastructure, supporting all types of service.

Network evolution — the interdependence of SDH and ATM

No two networks are alike, and network operators' priorities will be dictated by their own specific market situations, and whether they are established operators (i.e. defending monopoly or near-monopoly positions) or new operators (i.e. attacking existing operators). The following is a generic evolution that has aspects common to most network operators.

SDH deployment

The next few years will see large-scale deployment of SDH, in both existing and new public networks. SDH provides great flexibility in the network and also has considerable cost advantages in terms of operation and maintenance. Leased lines can be delivered with very short lead-times; the performance of the network can be tightly controlled and its utilisation maximised.

New operators can therefore be expected to build their networks using SDH from the start, and existing operators will migrate towards SDH as quickly as possible, in order to compete and offer new services.

In the inter-office network, SDH deployment will be driven by the need to increase capacity and manage the network. Major business centres could be interconnected early by SDH, thus providing end-to-end paths that can be used for ATM.

In the local network, SDH will first be deployed in those parts of the network that serve a high proportion of business customers. Ring configurations will be used to provide high speed and secure access. It is in these business areas that we believe ATM, too, will first be used.

Fig. 7 shows the layered architecture of the transport network, based on SDH.

ATM as an overlay network

The SDH network will provide the foundation for large-scale introduction of ATM. However, we believe the first ATM equipment will be deployed more or less in parallel with the first SDH systems. An overlay ATM network, based on both PDH and SDH, will be built, primarily to support data communications services.

SDH is optimised for Constant-Bit-Rate (CBR) services. When ATM is introduced, it will bring to the network an increased granularity and support of Variable-Bit-Rate (VBR) services. In general, the network will benefit from service independence.

As noted above, ATM will first be deployed in customer premises networks for high-speed LAN applications. Wide-area interconnection of CPNs will be achieved

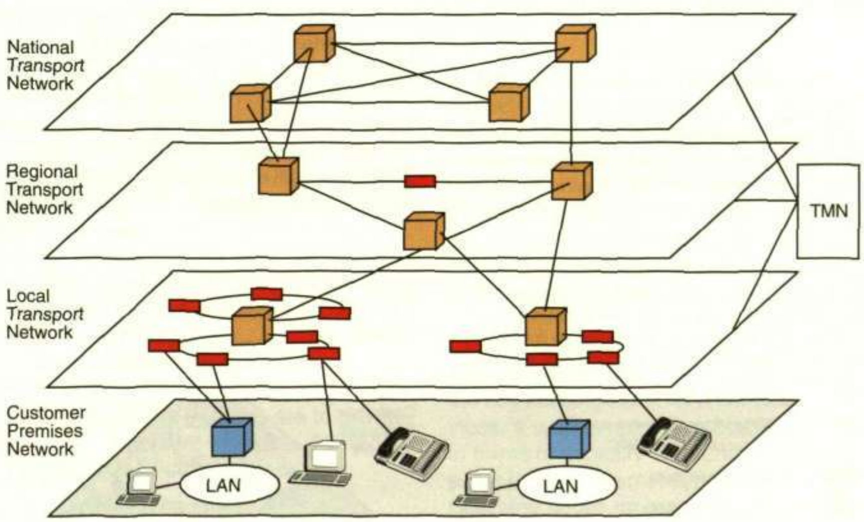


Fig. 7
A generic example of a layered SDH-based transport network. SDH will be deployed to provide the flexibility and performance needed to handle existing traffic. ATM will first be introduced in the CPN using the SDH leased-line service for interconnection between different locations

TMN — **Network management system, *e.g. according to TMN***

SDH DXC (cross-connect)

ATM Business Network Node

Add-drop multiplexer, ADM

using leased-line services over SDH or PDH networks.

As the use of ATM increases in the CPN, ATM can be introduced in the public network as an overlay network providing additional services, e.g. the VLL service, Fig. 8. SDH rings can be used for subscriber access, with a capacity of up to a full STM-1 (155 Mbit/s), the standardised access for B-ISDN. Existing traffic such as POTS can also be carried on this ring network, with remote multiplexers and other access nodes providing the final local-loop connection. The ATM access nodes can be shared for access to different services from one location and can include both data and voice, using different VP/VCs. In the ATM access node, the ATM traffic is concentrated to make more efficient use of the transport capacity. Other types of traffic are handled in SDH cross-connect systems.

The size of the ATM access node will vary, depending on the capacity required, from a small multiplexer to a large cross-connect. In the regional transport layer, ATM cross-connects can be used to route traffic between local areas. In the national transport network, transport is handled within SDH at the VC4 level, and ATM is not visible.

With the ATM overlay network in place, services such as Frame Relay and SMDS/CBDS can easily be added. As a first step, when demand is relatively low, these services can be introduced via a small number of central servers; as take-up increases, servers can be introduced in local access nodes.

Functionality for B-ISDN can also be added to both access and regional nodes by adding the appropriate hardware and software, Fig. 9.

ATM for narrowband services
As ATM technology matures and becomes more cost-competitive, it will also become economical to use ATM for narrowband services in the Public Switched Telephone Network (PSTN).

In a first step, ATM can be used for circuit emulation of 2 Mbit/s and perform transport and cross-connect functions between narrowband switching nodes. At a later stage, ATM could handle individual 64 kbit/s circuits and thus provide a ubiquitous infrastructure for switching and transport for all services.

Network management

To maximise the benefits of the SDH and ATM technologies, it is important to provide efficient management at network level from the very beginning. Management of SDH and ATM should be integrated so that routing, protection switching, performance monitoring and fault localisation are accomplished in an efficient way.

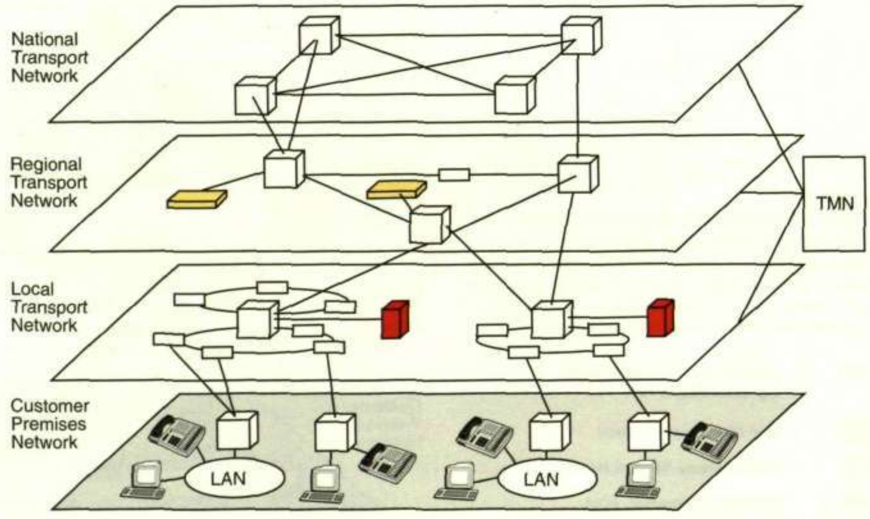


Fig. 8
An ATM network can be built by introducing ATM cross-connects in the transport network. The VLL service can be used to efficiently interconnect ATM-based CPNs. Other services, too, can easily be implemented

TMN — **Network management system, e.g. according to TMN**

ATM transit Node

ATM Access Node

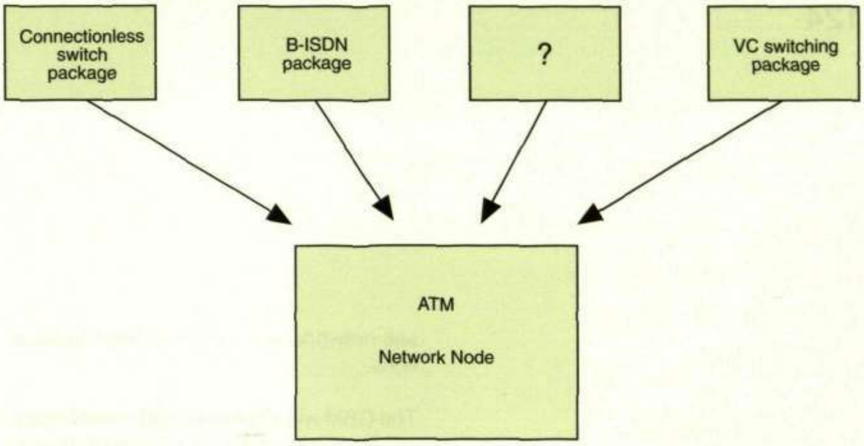


Fig. 9
Adding functionality to a node. The ATM node has to be scalable both with respect to capacity and functionality. New functions are added by supplying hardware and software packages

The required management functions can be divided into three interrelated parts:

Service management
Service management is concerned with the contracted services provided to customers or available to potential new customers. This includes different aspects of services and users as well as customer control. In implementation, service management will often be closely linked to other administrative systems and functions, e.g. order-getting and customer service.

Network management
This includes handling all the ATM network elements, both individually and as a set. At this level the management system is not concerned with how a particular element provides service internally. Both complete visibility of the whole network and vendor-independent views need to be maintained. Transport services at the virtual path and virtual channel levels are monitored.

Management of the ATM network will also include some new functions, such as routing algorithms for VBR connections, multiple QOS parameters, and congestion control.

Network element management
This includes handling of all network elements on an individual basis, providing a gateway to the network management level above it. It 'hides' the implementation details from higher levels of the management system, but maintains statistics, logs and other data relating to individual network elements.

Here too, new functions are included, such as the handling of individual QOS parameters for connections, variable bandwidth and congestion control.

The management systems are to provide all the management functions defined by the CCITT TMN standards, i.e.:

- Configuration management
- Fault management
- Performance management
- Security management
- Accounting management.

The Ericsson approach to broadband

Ericsson's strategy for the migration to broadband and multimedia communications is based on three core platforms:

- A full range of SDH systems, providing a flexible and fully-managed physical transport layer for all telecommunications services
- A generic, ATM-based platform – for broadband and multimedia services – that can be implemented in both public and private networks
- A fully-integrated management platform – for networks and services – that can provide management facilities end-to-end across both public and private networks.

These platforms, which can be combined to support different migration strategies, will be equally at home with new operators who are starting from scratch. In all cases, the emphasis is to support network implementations that generate revenue from the word go.

ETNA – a full line of SDH products
ETNA[3] (Ericsson Transport Network Architecture) encompasses a full range of SDH cross-connect and multiplexing systems, fully conforming to CCITT TMN standards. ETNA systems are commercially available, and are being installed by a number of public network operators in Europe.

The Generic Broadband Module
The Generic Broadband Module, GBM, is the key building block for ATM-based services. It is a completely scalable system, in terms of capacity and functionality: the same basic components are used in ATM switching nodes for both customer premi-

ses networks and public network applications.

The GBM will allow networks to evolve gracefully from an overlay ATM network, providing VLL functionality, to a high-capacity ATM infrastructure network providing broadband switching and B-ISDN.

The GBM will enter its first field trial with Deutsche Bundespost Telekom in Germany in 1994, where it will interact with ETNA equipment already installed.

Network Management — FMAS and the TMOS family
Network management within the ETNA family of SDH systems is provided by FMAS (Facility Management System), which provides a complete management environment for multi-vendor SDH networks, as well as for existing PDH systems.

FMAS is itself part of Ericsson's TMOS (Telecommunications Management and Operations Support) family of network management systems, which covers the full range of telecommunications services, including POTS, Intelligent Networks, mobile networks and business services. TMOS implements the CCITT TMN standards for network management.

Management of Ericsson's ATM systems will be based on the same TMOS platform. It will encompass end-to-end management of broadband services, including customer control of services, and provide full integration with management of the physical SDH transport layer.

Conclusion

The introduction of ATM and broadband services represents an evolutionary change in telecom networks. This change will be shaped – more than any previous one – by an increasingly dynamic and competitive commercial environment.

The initial deployment of ATM will be driven by the need to provide an infrastructure for competitive services. In a later phase, ATM will be used to provide a basic transport and switching infrastructure for all types of service.

Ericsson is currently completing a range of products and systems to support a flexible and operator-adapted introduction of ATM and broadband services.

References

1. Staxén P., and Vestin C-G.: *The Telecom Evolution in the Broadband Era.* Ericsson Review *70* (1993):1, pp 2-11.
2. Larsson M., Ljungberg M., and Rooth J.: *The ATM Switch Concept and the ATM Pipe Switch* Ericsson Review *70* (1993):1, pp 13-20.
3. Danielsson S.: *An Introduction to the Ericsson Transport Network Architecture.* Ericsson Review *69* (1992):3, pp 58-61.

ERICSSON

ISSN 0014-0171

Telefonaktiebolaget LM Ericsson

93428 Ljungföretagen, Örebro 1993

ERICSSON REVIEW

4 1993

Thermal Dimensioning of Air-Cooled Printed Board Assemblies
New Generation of MINI-LINK
Universal Personal Telecommunication (UPT) –
Concept and Standardisation
Evolving an Intelligent Architecture for Personal Telecommunication

ERICSSON REVIEW

Vol. 70, 1993

Contents 1993

ERICSSON REVIEW

Number 4 · 1993 · Volume 70

Responsible publisher *Bo Hedfors*

Editor *Per-Olof Thyselius*

Editorial staff *Martti Viitaniemi*

Subscription *Eva Karlstein*

Subscription one year USD 30

Address S-126 25 Stockholm, Sweden

Fax +46 8 681 2710

Published in English with four issues per year

Contents

Cover
Roof-mounted MINI-LINK in Gothenburg's harbour, Sweden. MINI-LINK is the generic name of an Ericsson family of compact radio links for point-to-point transmission of one or more 2 Mbit/s channels over distances up to 30 km per hop

Thermal Dimensioning of Air-Cooled Printed Board Assemblies

Åke Mälhammar

Modern electronic equipment is characterised by considerable power dissipation and therefore requires careful thermal dimensioning in order to ensure reliable function and long service life. The development of thermal dimensioning methods for electronic equipment has been rapid, and a new discipline of specialists has emerged in this field.

The author describes some models for heat dissipation on printed board assemblies and exemplifies some common methods of dimensioning air-cooled boards.

The rapid development in electronics has continuously reduced the volume and power required to implement a specific application. Since the reduction in volume has been greater than that of power, there has been a constant increase in power density, resulting in more and more hard-to-solve cooling problems, Fig. 1. There is no doubt that this trend will continue, although the extended use of opto-electronic equipment may eventually break it.

Cooling fluids

Air is by far the simplest and cheapest cooling fluid but, unfortunately, its effect is rather poor. However, this disadvantage must always be weighed against the problems and costs ensuing from the use of alternative cooling fluids. In fact, if air is compared with other fluids it is usually found to be the best choice. A marked trend in modern electronics is therefore to use air-cooling for printed board assemblies (PBA) with increasingly high power density. In many cases this has resulted in air-cooling devices of rather odd design.

An alternative to air cooling is liquid cooling. Small, enclosed two-phase circuits – so-called heat pipes – are used for cooling components in both telecom equipment and domestic electronic equipment. So far, large liquid-cooling systems have only been used in applications with very extreme demands on performance. Today, we can only speculate on whether such systems will be used more widely.

Dimensioning problems with air-cooled PBAs

A prerequisite for proper thermal dimensioning is that the basic physical model for the cooling process is correctly designed. A very common misconception in this con-

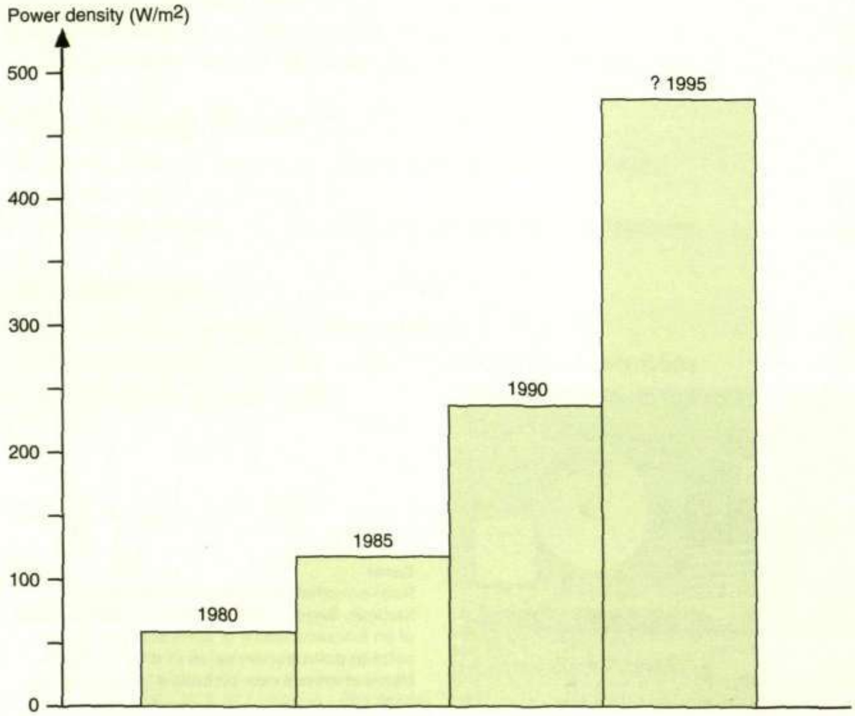


Fig. 1
The power density of printed board assemblies has been constantly increasing during the last decade. The values shown represent power losses in PBAs frequently used in telecom plants. There is no doubt that this trend will continue for some years to come but, in a longer perspective, the use of new technology may break it

ÅKE MÅLHAMMAR
Ericsson Telecom AB

Some common concepts used in thermal dimensioning.

Chip temperature	The temperature of a chip.
Cooling efficiency	A measure of the amount of power that can be removed from a PBA by cooling, in relation to an ideal reference case.
Forced convection	Air flow produced by mechanical means, such as fans.
Isothermal temperature	Constant temperature.
Natural convection	Air flow caused by rising air.
Pin temperature	The mean temperature at the point where a component terminal makes contact with the surface of a PBA.
Power density	Power per unit of area or unit of volume.
Temperature criterion	A measurement point and the corresponding temperature limit.
Thermal efficiency	A measure of the temperature distribution over the surface of a PBA.
Thermal territory	That part of a PBA which a component needs to be properly cooled.

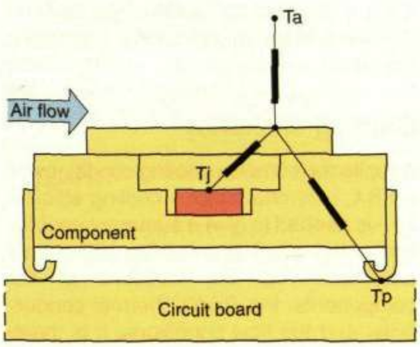


Fig. 2
Ericsson's thermal resistance network for components. The advantages of this model is that it is simple, that its constituent resistances can be measured, and that the accuracy of the model is high

T_p	**Pin temperature**
T_j	**Chip temperature**
T_a	**Air temperature**

text is that, on air-cooled PBAs, power is mainly dissipated by the surfaces of the components. In fact, the copper content of modern PBAs is so high that their thermal conductivity is fully comparable to that of the radiator fins of a motor car. When designing a model for thermal dimensioning of PBAs, it is therefore appropriate to regard the PBA as a large cooling fin on which a number of heat sources have been placed.

The portion of a component's generated heat that is dissipated through the PBA may vary greatly. For components the size of a thumb-nail and mounted on a PBA with a high copper content, it may be as high as 70 to 85 per cent. For components the size of the palm of the hand and mounted on a PBA with a low copper content, the corresponding figure may be as low as 10 per cent.

Since a large number of parameters affect the cooling of a PBA, thermal dimensioning is a complex problem. To be manageable it must therefore be split up into a number of subproblems. A natural division is to separate the components, the PBA surface and the cooling air.

The heat dissipation characteristics of a component are usually described in the form of a thermal resistance network. A large number of models for such resistance networks have been designed, the main difference being the number of resistances. The whole range from very simple networks containing only one resistance, to complex networks with as many as seven resistances is used today. Ericsson uses a three-resistance model, Fig. 2. The advantages of this model is that it is simple; its constituent resistances can be measured, and its accuracy is quite sufficient to meet the requirements of today's applications.

A resistance network is normally used to describe the thermal characteristics of the surface of a PBA as well. Since the number and locations of the components vary from one PBA to another, a resistance network prepared in advance cannot be used. Instead, this is done by programs which calculate the thermal properties of a PBA. These programs divide the PBA surface into small sections and use the pattern thus formed when designing the thermal resistance network, Fig. 3. Finite difference methods or finite element methods are used for this purpose.

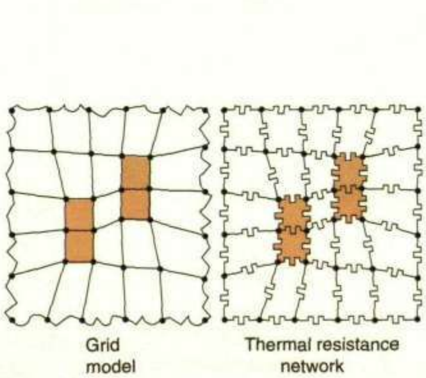


Fig. 3
Part of a PBA surface simulated as a thermal resistance network. The division of the surface into small sections is typical of the finite element method which, unlike the finite differential method, permits the use of non-rectangular cells

Characterising the properties of cooling air is a difficult task. The flow problems are usually complex and the surfaces of the PBAs are also uneven. There are two basic ways of solving the flow problem: classical methods and numerical methods. The numerical methods often give better results but they also require much larger calculation capacity. Since the cooling-air problem is combinative (flow and heat dissipation) it cannot be characterised by

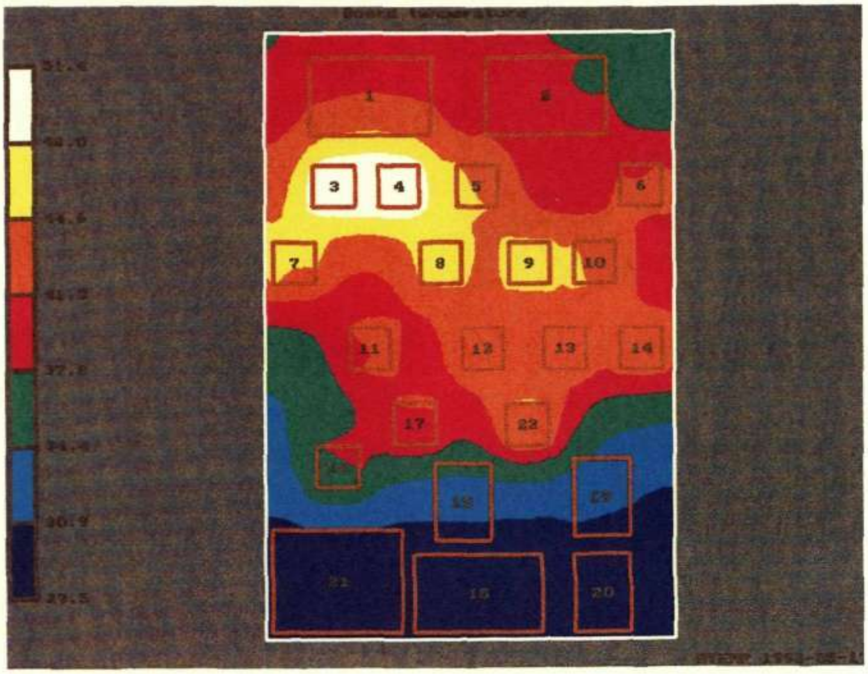


Fig. 4
Temperature distribution on a PBA. Each colour *corresponds to a given temperature interval.* This type of calculation has become a standard procedure

means of a resistance network. However, it is always possible to create a model which simulates the cooling air with equivalent thermal resistances.

Converted into numerical form the PBA cooling problem gives an equation system with a few thousands unknown quantities, which, of course, is a simple task for today's powerful computers. The result of the calculation is often very illustrative, Fig. 4.

Cooling efficiency

Despite the complex cooling conditions on a PBA, only one factor – cooling efficiency – is needed to give a summary description of the effect of a number of parameters, such as the locations of the components, the PBA's thermal conductivity, and the flow conditions. It is therefore possible to use the following simple formula to calculate the cooling effect on a PBA:

$$P_{real} = \eta_c \cdot P_{ideal} \qquad (1)$$

where
P_{real} = heat dissipation for the real case
η_c = cooling efficiency
P_{ideal} = heat dissipation for the ideal case

The ideal case is defined as two parallel plates whose isothermal temperature is at the same level as the maximum temperature in the real case, Fig. 5. Since the ideal case is simple and well defined, it is also relatively easy to calculate. Most PBAs have a cooling efficiency of less than 100 per cent, but for PBAs with very large surface extensions (such as cooling fins) the cooling efficiency may exceed 100 per cent.

Any measures which will equalise the temperature distribution on a PBA, and thereby approximate the ideal case, will increase the cooling efficiency. In practice this can be achieved in two ways: by equalising the power distribution or by increasing the thermal conductivity of the PBA. Fig. 7 shows the cooling efficiency levels

Fig. 5
The definition of cooling efficiency for a PBA, defined as the quotient between the cooling effect of the real case and the cooling effect of an ideal case. The highest cooling efficiencies are achieved for PBAs with high thermal conductivity and uniform power distribution. The *cooling-efficiency concept may well be used for* making thermal estimates at an early stage of the design process

$\eta_c = P_{real}/P_{ideal}$

P_{real} **Heat dissipation for the real case**
P_{ideal} **Heat dissipation for the ideal case**
η_c **Cooling efficiency**
T_a **Air temperature**
T_{bmax} **Max board temperature**

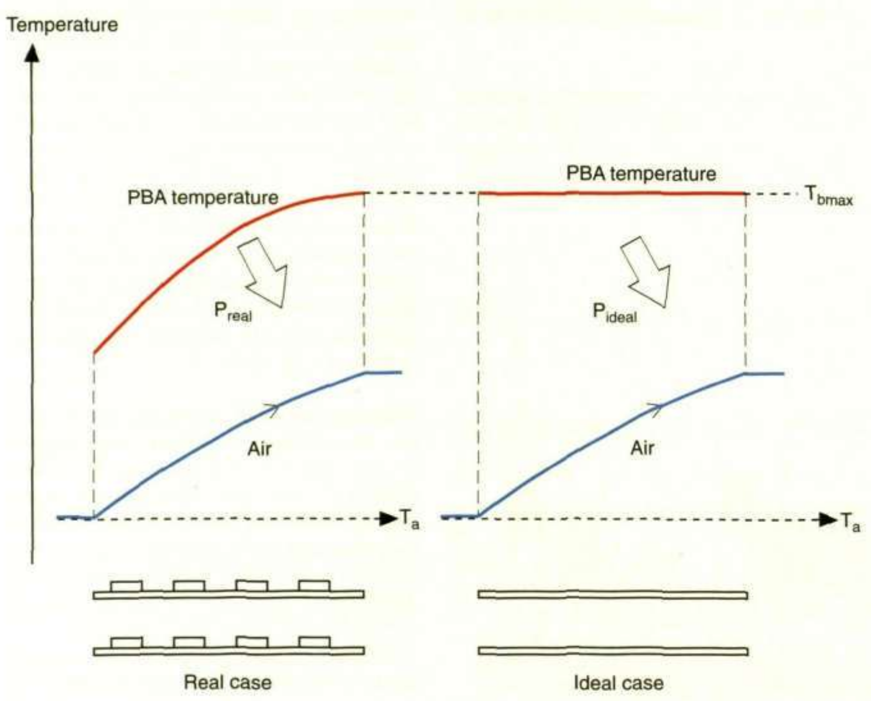


Fig. 6
Ericsson Telecom has a well equipped thermal laboratory. Measurements can be made in the whole range, from chip to system level

that can be reached with careful thermal dimensioning and the PBA technology used by Ericsson today.

Front-end and back-end thermal calculation methods

A number of important choices must be made at an early stage of the design process. For these choices to be as good as possible, the thermal consequences of different alternatives must be analysed, and this requires simple and rapid methods suitable for calculation at this early stage. In later phases of the design work, high-precision methods will be required.

At an early stage of the design process, a number of important parameter values are subject to great uncertainty. For this reason, calculation with the input data available at that time cannot be expected to provide accurate results, although they suffice as a basis for making rapid comparisons between different design alternatives. There are a number of methods of front-end thermal dimensioning. Ericsson normally uses both a PBA-oriented method (the cooling-efficiency method) and a component-oriented method (the thermal-territory method).

A program for thermal calculation of PBAs is normally used for thermal dimensioning at later stages of the design process. Basically, this type of dimensioning means grouping the components on the PBA so as to provide the best possible temperature conditions. A couple of years ago it was still possible to achieve a considerable increase in cooling efficiency by this method. Improvement figures of up to 20 per cent were not unusual. In later years, however, regrouping the components has become very restricted for electrical reasons, and thermal improvements have proven less successful.

The common use of high signal frequencies requires narrow spaces between components, and the large number of connections on the PBAs reduces the number of alternative layouts. This accentuates the need for methods that can be used in the systems design phase. The prime purpose of back-end calculation today is to verify that the cooling is satisfactory and to provide working data for calculation of the expected fault intensity of the PBA.

The cooling-efficiency method

With this method cooling efficiency is estimated and the result is then used in a relatively simple calculation according to Equation 1. Although the method is theoretically correct, the accuracy of the result can never be better than the accuracy of the estimate. However, the uncertainty of the method must be weighed against the uncertainty of the input data. For some properties, such as the power dissipation of components, this uncertainty may correspond to a factor of 2 at an early stage of the design process. Cooling efficiency can be estimated on the basis of empirical data of the type illustrated in Fig. 7.

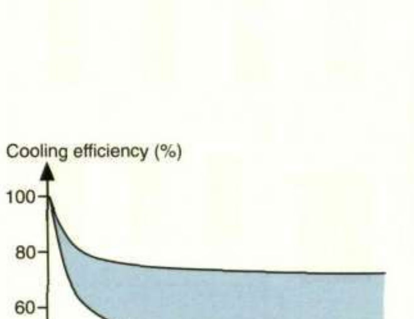


Fig. 7
Cooling-efficiency values that can be obtained by careful thermal design and use of Ericsson technology. The diagram applies on condition that the chief cooling method used is convection

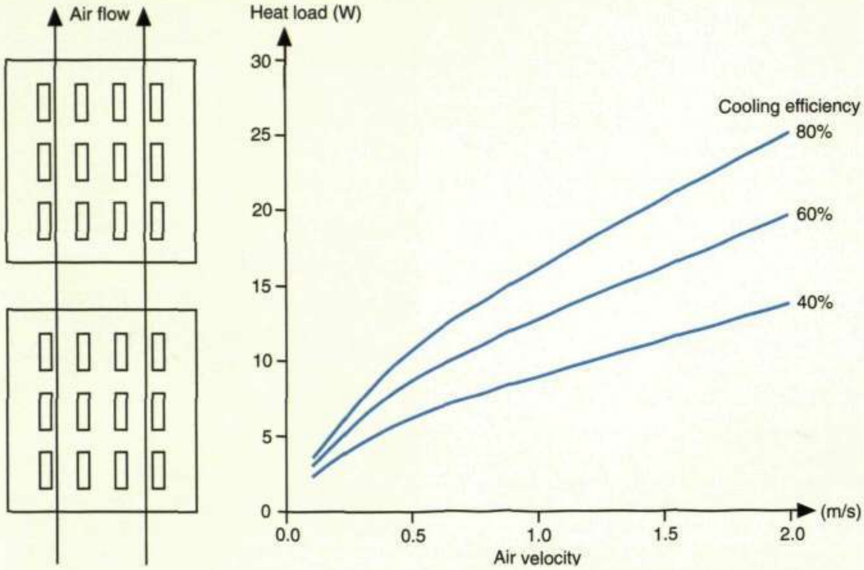


Fig. 8
Example of a thermal estimate by means of the cooling-efficiency concept. The cooling air enters the rack at the bottom and is heated on its way upwards, succesively degrading the cooling power of the air. This type of calculation belongs to the front-end design methods and is particularly useful at the very early stages of the design process

Size	200 · 200 mm
Pitch	16 mm
Temp. criterion	$\Delta T=25^{\circ}C$

When the cooling-efficiency method is put to practical use, calculations are based on the highest permissible temperature on the PBA, and then the ideal case and the cooling efficiency are used to calculate the power dissipation that the PBA can tolerate. Fig. 8 shows a typical result. The method is PBA-oriented and does not give any direct information about internal component temperatures. For this reason, calculations of cooling efficiency are usually supplemented with one or more simplified thermal calculations for the constituent components.

The thermal-territory method

This method is based on the fact that each component requires a certain portion of the area of the PBA to be properly cooled: what is called the thermal territory. The approximate thermal territory of an individual component can be calculated by replacing the component with an equivalent, round heat source, and then assuming that the heat will spread equally in all directions. The problem encountered here is equivalent to the problem of finding the diameter of a circular cooling fin that will cool a given power at a given difference in temperature. This problem is easy to solve by numerical methods.

In practice, handling circular surfaces is difficult, and Ericsson therefore uses a method which calculates rectangular thermal territories in a similar manner. The result of the calculation can be used directly for preliminary location of components, Fig. 9. By combining the thermal-territory model with a thermal-component model, the fact that part of the heating effect is emitted directly by the component surface is also taken into account.

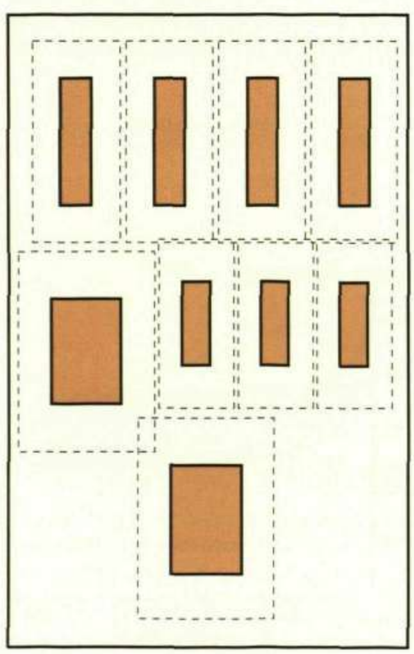


Fig. 9
Use of the thermal-territory concept for a PBA layout. If the territory surfaces overlap at some point, the respective components must be capable of expanding the surface in some other direction

Fig. 10 shows a typical temperature profile for a thermal territory. The temperature of the territory decreases with the distance from the component, which means that the cooling provided by the outer surfaces is not as efficient as that provided by the surfaces closest to the component. Thermal efficiency is a measure of the efficiency of the thermal territory and is defined as the quotient between the mean excess and maximum excess temperatures of the thermal territory. Too large thermal territories make poor use of the cooling provided by the PBA surface, resulting in greatly reduced thermal efficiency, Fig. 11. The rule of thumb is that the thermal efficiency of a component is allowed to fall below 75 per cent only in exceptional cases, and that the value must never be below 50 per cent.

For a PBA without surface extensions, the thermal efficiency for the PBA's components may be used as basic data when

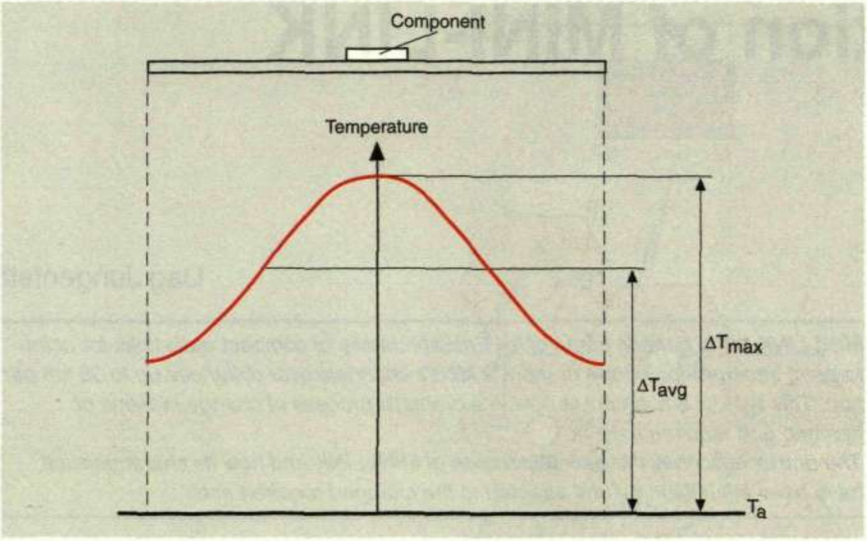


Fig. 10
Typical temperature profile for a thermal territory. The thermal efficiency is obtained by a comparison between the mean excess temperature of the territory surface and the maximum excess temperature of that surface. The rule of thumb is that the thermal efficiency must be at least 75 %

$\eta_t = \Delta T_{avg} / \Delta T_{max}$

T_a	*Air temperature*
ΔT_{avg}	**Mean excess temperature of the territory surface**
ΔT_{max}	**Max excess temperature of the territory surface**
η_t	**Thermal efficiency**

estimating the cooling efficiency of the PBA as a whole. Normally, in this case cooling efficiency is approximately equal to the lowest thermal-efficiency value.

However, the heat flux conditions on a real PBA are not as favourable as those assumed to apply when calculations according to the thermal-territory method are made. Consequently, to prevent the maximum temperature on a PBA from exceeding the set temperature limit, the total area of all thermal territories must be smaller than the area of the PBA. The rule of thumb is that the thermal territories must not take up more than 50 to 75 per cent of the area of the PBA.

Conclusions

The power density on PBAs has increased continuously. *To prevent quality problems* caused by high component temperatures, proper thermal dimensioning is therefore a necessity.

The problem of dimensioning air-cooled PBAs is complex, but calculation methods are provided which give satisfactory results even early in the design process. The development in PBA technology in recent years has emphasised the importance of front-end thermal design methods.

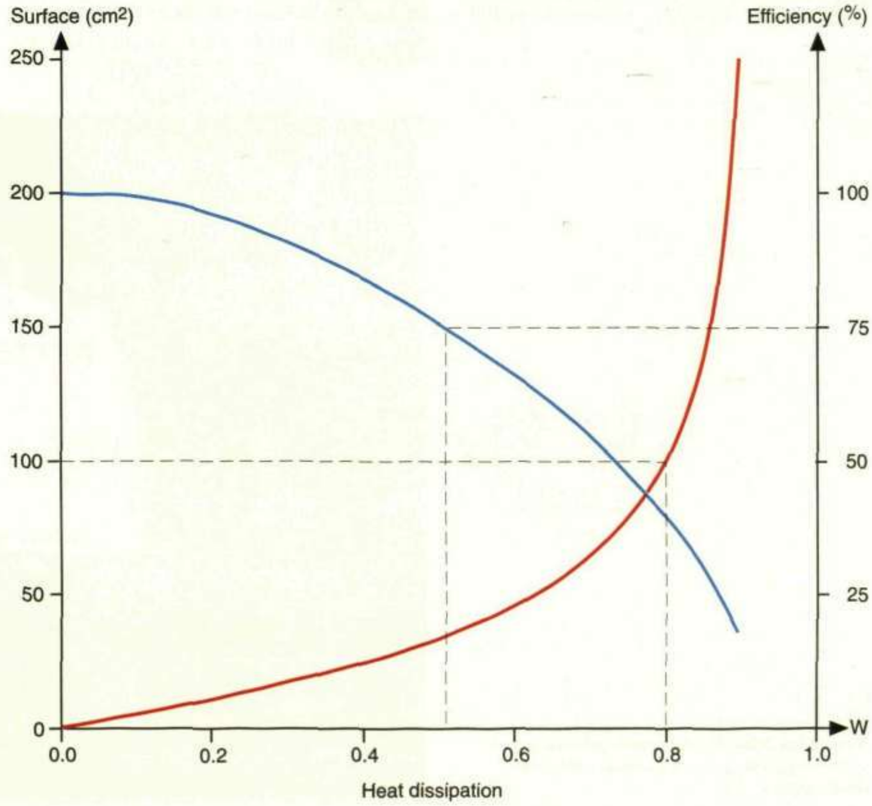
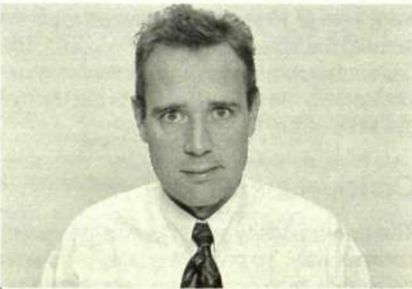


Fig. 11
Calculation of the variation of the territory surface and thermal efficiency with the heat dissiapation. If the above-mentioned rule is applied (thermal efficiency must be at least 75 %), then the component concerned must not be loaded with more than 0.5 W. If this limit is exceeded (the component loaded with, say, 0.8 W) the territory surface will be 100 cm² **(10·10 cm), which in most cases is an unreasonably large surface**

New Generation of MINI-LINK

Dag Jungenfelt



DAG JUNGENFELT
Ericsson Radar Electronics AB

MINI-LINK is the generic name of an Ericsson family of compact radio links for point-to-point transmission of one or more 2 Mbit/s channels over distances up to 30 km per hop. This type of equipment is now in a dramatic process of change in terms of demand and requirements.
The author describes the new generation of MINI-LINK and how its characteristics have been influenced by and adapted to the changed requirements.

Since MINI-LINK was introduced in the late 70s, some 20,000 units have been delivered and put into service in more than 70 countries. Having a transmission capacity of between 2 and 8·2 Mbit/s and a range of up to 30 km, MINI-LINK is often a cost-effective alternative to wired transmission systems in
- mobile telephone networks, Fig. 2
- public networks using AXE systems, Fig. 4
- private networks using MD 110 systems, Fig. 3.

The accelerated sales volume of MINI-LINK in the last few years is mainly attributable to the large investments in mobile telephone networks all over the world. Our industry's focus on a rapid expansion of the telecom networks in Eastern Europe and other regions has also increased customer interest in MINI-LINK. There is every indication that this widespread acceptance of MINI-LINK will continue, Fig. 5.

The introduction of a new generation of MINI-LINK is therefore a natural step in Ericsson's product development programme, in response to new requirements from a growing market and based on experience from the company's 15 years of operations in compact microwave links.

New demands

The establishment of new mobile telephone systems with smaller cells and higher traffic-handling capacity implies an increased number of base stations and, hence stresses the need for effective

Fig.1
A compact MINILINK is an efficient method of connecting a downtown sales office to the headquarters

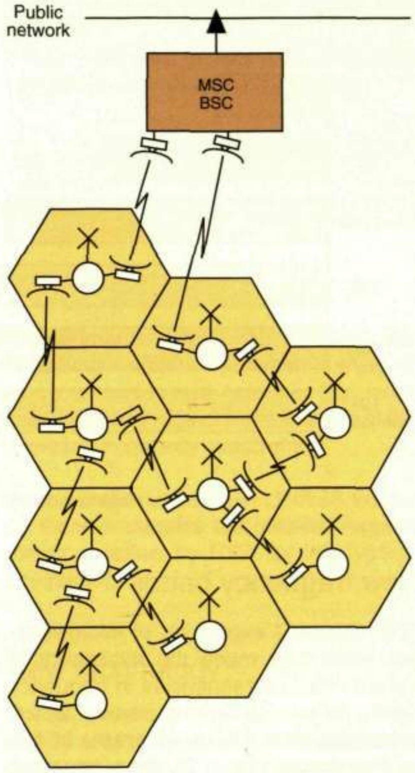


Fig. 2
Microwave link technology provides efficient means of connecting radio base stations to switching centres

MINI-LINK
Base station

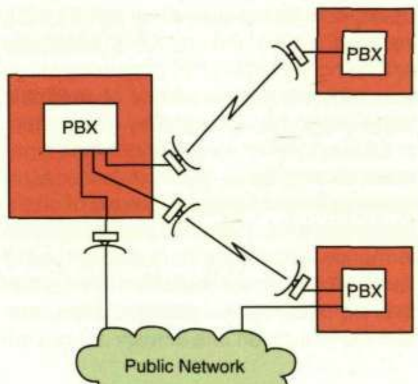


Fig. 3
For private networks, microwave links are often a good solution when interconnecting major nodes

transmission. Often, a microwave-link configuration offers the quickest and most cost-effective method of setting up such connections. See the comparison of costs in Fig. 6.

The ongoing development has changed the market for MINI-LINK. Starting as a product with its application limited to isolated point-to-point connections, MINI-LINK has become a volume product used as a building block in complex transmission networks. This accentuates the demand for a MINI-LINK design that can be integrated in the superordinate system – *functionally, mechanically and electrically.*

For this reason, the new generation of MINI-LINK is designed so as to allow integration into other Ericsson systems, in particular telecommunication management systems. The design of MINI-LINK products has also been adapted to large-scale production and installation.

MINI-LINK is a general product for connections carrying n·2 Mbit/s in different types of public and private networks, and primarily designed to meet increased demands for more effective transmission systems in mobile telephone networks.

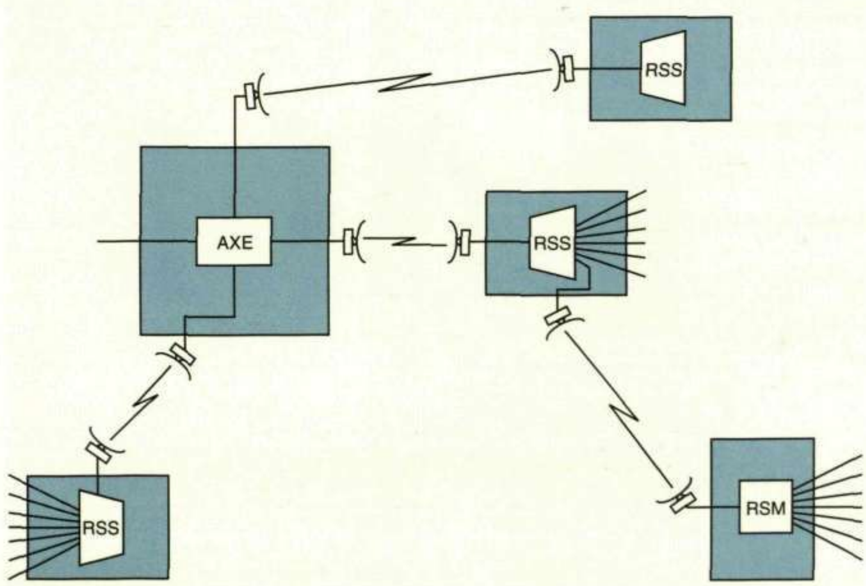


Fig. 4
Remote subscriber access units are conveniently connected to the network by microwave links

RSS Remote subscriber switch
RSM Remote subscriber multiplexer

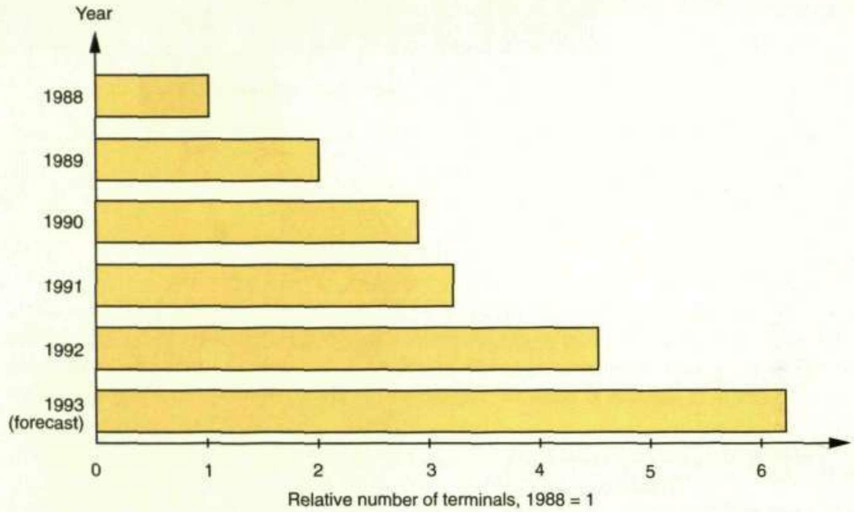


Fig. 5
MINI-LINK sales.
The increased use of MINI-LINKs in cellular networks has promoted MINI-LINK sales

New frequency bands

The increased expansion of microwave-link traffic has made the availability of vacant channel frequencies in frequency bands below 20 GHz a limiting factor, especially in metropolitan areas. In frequency bands above 20 GHz, however, channel frequencies reserved for radio link applications are in good supply.

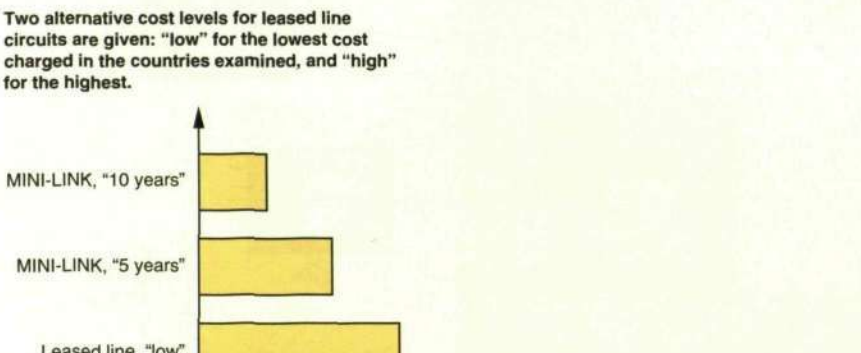
Up till now, two thirds of all MINI-LINKs installed are in the 15 GHZ band. By extending the MINI-LINK frequency range up to 38 GHz, the number of available frequencies are increased by a factor ten. In addition higher frequency means higher directivity for a given antenna size, resulting in more effective re-use of channel frequencies. Improved equipment performance also means more efficient use of spectrum together with better tools for frequency planning. In practice, the drawbacks – in the form of limited availability of channel frequencies – have been eliminated, and the possibility of adapting frequency selection and output power to local, regulatory conditions has improved considerably.

Fig. 6
MINI-LINK vs leased circuits – a comparison of costs

The figure shows the relative, annual transmission cost in a cellular network. The comparison is based on the need for transmission of a typical segment in a GSM network, the total cost of establishing and operating a MINI-LINK network, and the cost of leased lines in Europe.

The depreciation period for MINI-LINK is assumed to be five and ten years respectively.

Two alternative cost levels for leased line circuits are given: "low" for the lowest cost charged in the countries examined, and "high" for the highest.



New requirements

The new fields for application of MINI-LINK introduces more demanding functional requirements as well as intensified requirements for operational performance. The transmission quality in terms of acceptable bit error ratio, availability, etc, must be improved, and so must the spectral characteristics in order to permit effective utilisation of the available bandwidth. The scope of requirements in the form of directives, standards and recommendations issued by national, regional (CEN/CENELEC, ETSI) and global (ITU, IEC) organisations is constantly widening.

The MINI-LINK now being introduced meets these requirements. The following characteristics, to name a few, have been improved as compared with previous MINI-LINK versions:

- The transmitters feature narrower bandwidth, extremely low levels of unwanted output signals, and continuously adjustable output power
- The receivers feature improved detection performance and more effective suppression of unwanted signals
- The choice of antennas has been supplemented with high-performance versions, which reduce the radiated power in unwanted directions by 20 dB; and with dual-polarised versions that double the capacity over a route
- EMC properties have been improved and surpass the requirements specified in existing standards.

The enhanced receiver detection performance thus achieved makes it possible to considerably improve the transmission quality on individual routes. All other

Table 1
A comparison of size and weight between the new MINI-LINK and its predecessor (without antenna)

	New generation	Old generation
Number of units	2	5
Weight	8 kg	22 kg
Dimensions	420·330·230 mm	650·370·390 mm

improvements serve to reduce the disturbing effects between adjacent systems, and to use narrower frequency spacing to ensure more efficient use of the available frequency spectrum.

Performance data of MINI-LINK 23, 26 and 38 meet or surpass the detailed requirements specified by ETSI for this type of equipment.

New technology

The new versions of MINI-LINK 15, 23, 26 and 38 are the result of intensive technical development work based on know-how acquired through 15 years of experience in compact microwave links. The technical design of the new generation of MINI-LINK has been guided by the following key concepts:

Integration, for low weight, small volume and ease of handling in all phases
Automation, for high quality and short throughput times in all phases.

While the number of functions has increased and performance has improved, the weight, volume and number of constituent physical units of MINI-LINK have decreased drastically; Table 1. This has been achieved by consistent integration and a completely new packaging practice for the microwave parts. Both these factors have clearly contributed to greater dependability.

The documented MTBF value for the older version of MINI-LINK – 18 years per terminal – is therefore expected to be considerably greater for the new generation of MINI-LINK.

All active microwave circuits are accommodated in a single microwave unit, and the proportion of ASICs and MMICs (monolithic microwave integrated circuits) has increased. This, together with new, commercially available microwave components (FETs = field effect transistors and HEMTs = high electron mobility transistors) and the new packaging practice, has contributed to:
- Enhanced performance through the use of better components and reduced losses in internal transitions
- Shorter throughput times and better quality through advanced automation of mounting, wiring and testing.

It has also been possible to reduce the number of versions by applying solutions

Fig. 7
Installation test of a MINI-LINK mounted on the mast for a cellular phone, radio base station antenna placed on the roof of an industrial building in Gothenburg, Sweden

based on broader bandwidths than those used in previous versions.

The new packaging practice has required careful modelling and characterisation of all constituent subsystems. This design work has generated a large library containing detailed computer models of amplifiers, filters, oscillators, etc., which provide a valuable basis for the continuing development and improvement work.

System characteristics

If a radio-relay system is to serve as an efficient element in the network, high radio performance and flexibility, as regards capacity and channel frequencies, are not the only prerequisites. Effective operations support and well-adapted accessories and tools are integral parts of the MINI-LINK concept.

System configurations
The basic version of MINI-LINK has a transmission capacity of 2, 2·2, 4·2, 8·2, 8, 2·8, or 8 + 4·2 Mbit/s. Division into tributary channels is performed by a secondary multiplexer in the indoor part of the link. The introduction of versions based on US and Japanese standards is planned for 1994.

Today, MINI-LINK is produced for operation in the following frequency bands:

- 14.50 – 15.35 GHz MINI-LINK 15
- 21.20 – 23.60 GHz MINI-LINK 23
- 24.25 – 29.50 GHz MINI-LINK 26
- 37.00 – 39.50 GHz MINI-LINK 38

The choice of frequency band for a given project is determined by two factors: regulatory conditions and requirements for hop length and availability, Box A.

Redundant or protected radio link hops, 1+1 systems, can be implemented by employing frequency diversity, space diversity, or hardware redundancy exclusively. Two signals are processed in parallel on the receiving end and the output from one of the receivers is selected. If the signal received by the active receiver is deteriorated, the traffic is switched over to the standby receiver without loss of traffic (so-called hitless switching). MINI-LINK 23

Box A
Availability

Since microwaves are attenuated by rain, the dimensioning of high-frequency radio link connections must include a signal strength margin. This margin is determined by the following factors:
-the frequency band used
-the length of the connection
-the climatic zone (expected rainfall intensity)
-availability and quality requirements.

For a given set of parameters (frequency, distance and climate) we can calculate the probability of the quality of the connection falling below a given level. Under normal operating conditions, a MINI-LINK connection is practically without bit errors.

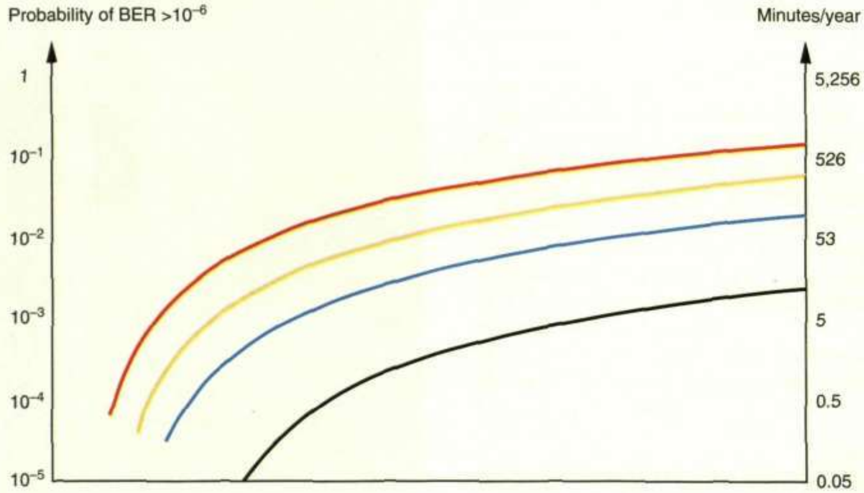


Fig. A
The diagram shows the probability of the bit error ratio (BER) exceeding 10^{-6} for MINI-LINK 15, 23, 26 and 38. The curves apply to areas where the rainfall intensity does not exceed 42 mm/hour for more than 0.01 % of the time

MINILINK 15
MINILINK 23
MINILINK 26
MINILINK 38

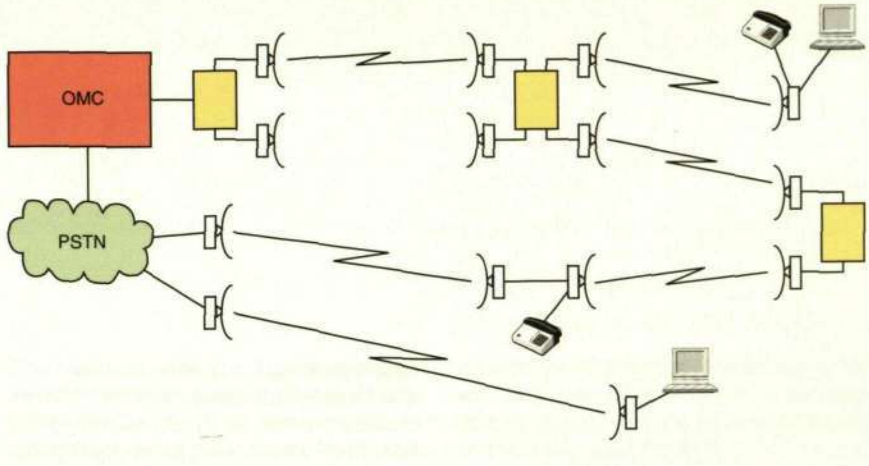


Fig. 8
MINI-LINK has a built-in, microprocessor-based control and supervisory system. Service phones can be connected to each MINI-LINK terminal. An OMC (operation and maintenance centre) terminal or a PC connected to any MINI-LINK terminal gives full access to the control and supervisory system of the MINI-LINK network

also incorporates 2+1 systems (two 8·2 Mbit/s terminals and a common standby terminal) for protected 16·2 Mbit/s hops.

Operations support

MINI-LINK has a built-in, microprocessor-based control and supervisory system whose main functions are:

- Performance supervision of the link connection (transmission quality and availability)
- Control of signal loops for function checking and fault tracing
- Alarm collection for status checking and fault tracing
- Supervision/control of user-defined inputs and outputs.

The interfaces and protocols of the supervisory system are compatible with all members of the MINI-LINK family, and all radio links in a network can be supervised from any network terminal in the network, Fig. 8.

All control and supervisory functions are accessed through a terminal interface for connection to a handheld terminal or a PC. The terminal interface also serves as a connection point for central operations support systems, such as MNM (MINI-LINK network manager) for PCs, or Ericsson's telecommunications management and operations support platform, TMOS, which includes a subsystem called microwave radio operation and maintenance system, MOMS.

Accessories and tools

In addition to radio terminals, antennas, operations support systems, multiplexers and switching units for 1+1 systems, the MINI-LINK family also includes a complete set of accessories and tools. The accessories programme has been developed in close collaboration with customers and users with a view to facilitating engineering, installation and operation of MINI-LINK systems. The following are examples of programme components:

- Performance prediction and calculation tools for route dimensioning and frequency planning
- Installation fittings in the form of tripods, wall brackets and pole brackets
- Units for adaptation to other Ericsson equipment, such as radio joint box (RJB), for smooth integration with AXE
- An interface with improved lightning protection between indoor and outdoor equipment

A common characteristic of all programme components is that they serve to widen the field of application of MINI-LINK and to reduce users' engineering, installation and operating costs.

System components

MINI-LINK consists of three main components or modules, Fig. 9:

- An outdoor radio module capable of performing all functions of a complete radio link terminal
- An antenna module which is either integrated with the radio module or installed adjacent to it
- An indoor access module for systems with four to eight traffic channels, or for systems with redundancy. For ease of installation, an access module may also be used with systems having only one or two traffic channels.

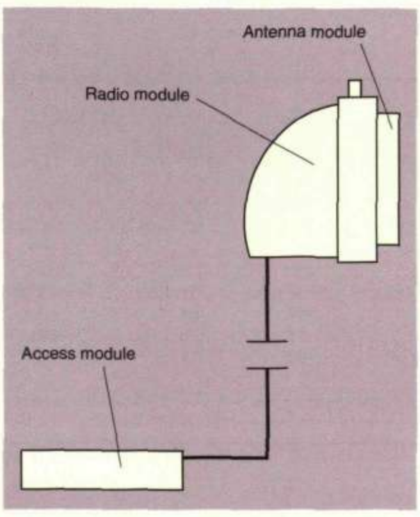


Fig. 9
Main components of MINI-LINK

Radio module
The radio module consists of two electronic units housed in a cast, weatherproof box which is protected against electromagnetic radiation. This box contains all MINI-LINK functions, except multiplexers and redundancy switches:
- Traffic interfaces: 2, 2·2, 8 or 2·8 Mbit/s according to CCITT's Rec. G.703
- Supervisory functions, including a terminal interface to a handheld terminal, PC or a central operations support system
- Analogue service channel for connection of a service telephone.

The interface to the antenna is a standard IEC waveguide. The radio module can be used in stand-alone mode without an access module and with any type of antenna.

Technical performance data of the radio module are shown in Box B. These data meet or exceed issued and preliminary standards for the frequency bands concerned. The transmitter's output power can be adjusted to values in the 0-50 dB range, which ensures efficient re-use of channel frequencies. The electronic equipment of the radio module consists of a baseband unit and a microwave unit.

The baseband unit contains all bitrate dependent parts but is independent of the frequency band used. The same baseband units can therefore be used for the whole of the new MINI-LINK programme.

The microwave unit contains all parts that are dependent on the frequency band used. The RF oscillators in MINI-LINK are controlled by a synthesiser. Channel frequencies can be selected in steps of 1.75 MHz within a 280 MHz wide band for MINI-LINK 38 and a 560 MHz wide band for MINI-LINK 23 and MINI-LINK 26.

The use of the same microwave units, regardless of bit rate, simplifies upgrading and spare parts service.

Antenna module
The antenna module consists of a parabolic antenna, a radome, a bracket and an interface to the radio module. Some versions have integrated mechanical components for the radio and antenna modules, Fig. 10, and some use the same bracket, Fig. 11. A common feature of all versions is that the radio module can easily be dismounted without affecting the pointing of the antenna.

Box B

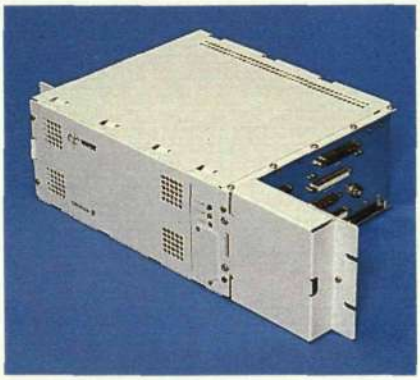
MINI-LINK, technical data

	MINI-LINK 15		MINI-LINK 23		MINI-LINK 26		MINILINK 38	
Frequency band	14.50-15.35 GHz		21.2 - 23.6 GHz		24.25 - 29.50 GHz		37.0 - 39.5 GHz	
Traffic-handling capacity	2, 2·2, 4·2, 8·2, 8, 2·8 or 8+4·2 Mbit/s							
Channel spacing	3.5, 7 & 14 MHz							
Output power	19 dBm (26 dBm opt.)		19 dBm		11 dBm(19 dBm opt.)		15 dBm	
Antenna size	60 cm	120 cm	30 cm	60 cm	30 cm	60 cm	30 cm	60 cm
Antenna gain,	37.0 dBi	42.5 dBi	34.5 dBi	40.5 dBi	35.5 dBi	41.5 dBi	38.0 dBi	44.0 dBi
Receiver threshold	10^{-3}	10^{-6}	10^{-3}	10^{-6}	10^{-3}	10^{-6}	10^{-3}	10^{-6}
2 Mbit/s	-93 dBm	-87 dBm	-92 dBm	-87 dBm	-91 dBm	-88 dBm	-88 dBm	-83 dBm
2·2	-86	-81	-86	-81	-85	-80	-85	-80
8	-83	-78	-83	-78	-82	-77	-82	-77
2·8	-80	-75	-80	-75	-79	-74	-79	-74
Traffic interface	CCITT Rec. G.703							
Terminal interface	RS 232 (V24/V28)							
Service channel	4-wire/600 ohm							
Power consumption	<30 W		<30 W		<30 W		<35 W	

Fig. 10
In some versions of MINILINK, the radome with the parabolic antenna and the radio module are contained in one mechanical unit

Fig. 11
The picture shows the radio module and antenna radome monted on a mast. A common feature of all MINI-LINK versions is that the radio module can easily be dismounted without affecting the pointing of the antenna

Fig. 12
The access module is available in two mechanical versions: for installation in a 19" rack and for wall mounting



The standard antennas for all frequency bands have a diameter of 30 and 60 cm. MINI-LINK 23 and 15 can also use 120 cm antennas.

The radomes protect the antennas against snow and ice and can also be equipped with attenuating material to furter limit wide-angle side lobes and back lobes.

For antennas that are not mechanically integrated with the radio module, the interface consists of a short, flexible waveguide.

All antennas have a mechanism for fine adjustment of the pointing in both azimuth and elevation.

Access module

The access module is MINI-LINK's interface to the network. The basic access module consists of an electromagnetically shielded unit, having a connection point for termination of the cable from the radio module, as well as a number of function-specific user interfaces. The access module can be equipped with one or two secondary multiplexers and/or a switching unit for redundant systems.

The access module is available in two mechanical versions: for installation in a 19" rack and for wall mounting. In a 19" rack, the basic version requires a building height of 44 mm (1U). If multiplexers or switching units are included, the required building height is 132 mm (3U), Fig. 12.

Summary

The present development of telecom markets – and, in particular, the expansion of mobile telephony – increases the demand for cost-effective transmission methods. Microwave links are a flexible tool, provided that capacity, frequency band, functionality, price, performance and quality meet specified requirements. Ericsson's new MINI-LINK generation is designed to meet these requirements.

References

1. P.O. Nilsson and S. Reinefjord: *MINI-LINK Mk II, a new member of the MINI-LINK family.* Ericsson Review *65* (1988):2, pp. 69-74.
2. ETSI Transmission and Multiplexing: *Parameters of radio relay systems for the transmission of digital signals and analogue video signals operating at 23 GHz,* pr ETS 300 198, January 1992.
3. ETSI Transmission and Multiplexing: *Parameters of radio relay systems for the transmission of digital signals and analogue video signals operating at 38 GHz,* pr ETS 300 197, January 1992.
4. ETSI Transmission and multiplexing: *Draft technical specification for fixed point to point radio link equipment for the transmission of digital signals and analogue video signals operating in the frequency range 24.250 to 29.500 GHz.* TM4 (93) 62, annex 1a.

Universal Personal Telecommunication (UPT) – Concept and Standardisation

Jonas Sundborg

Universal personal telecommunication , UPT, a new service concept in the field of telecommunications, aims at making telecommunications both universal and personal. Instead of calling a telephone line or a mobile terminal, you call the person you wish to get in touch with and leave it to the network to locate the line or terminal where he/she can be reached.

The author describes basic concepts behind UPT, the UPT service concept, standardisation activities and the evolution of UPT. Ericsson's views of the concept and products for the UPT service will be presented in the next issue of Ericsson Review.

Trends in telecommunications indicate an increasing demand for mobility of teleservices. People move more often, travel more rapidly and spend more time on work outside the traditional office environment. Current deregulation and liberalisation activities are also affecting the provision and use of teleservices.

This mobility can be either terminal mobility, as in cordless telephone and cellular mobile systems, or personal mobility as in universal personal telecommunication, UPT. UPT, which supports user identification instead of identification of a fixed line or a portable terminal, will efface the boundaries between fixed and mobile networks.

Concepts

Network access identification

In today's public switched telephone network, PSTN, the access port, AP (the point of attachment of the line), defines the subscriber's identity to which charges related to the use of the line can also be assigned. This is indicated by (1) in Fig. 2. Consequently, there is no need for identification of terminals – any appropriate terminal can be used, Fig. 3. Wired access can also be extended to comprise wireless terminals, e.g. the first generation of cordless telephones, CT1.

Fig. 1
Instead of calling a telephone line or a mobile terminal, you call the person you wish to get in touch with and leave it to the network to locate the line or terminal where he/she can be reached

JONAS SUNDBORG
Ericsson Telecom AB

Identification of terminals

A step further from network access identification is the supplementary service *terminal portability* supported by the integrated services digital network, ISDN. This supplementary service enables the calling and/or the called party – within the same basic access, during the active state of a call – to unplug their terminal from one socket and plug it into another, Fig. 3.

Mobile telecommunication networks are based on *terminal mobility*. This means that the terminal is allowed to roam within the area covered by the mobile network. The point of attachment will change each time the terminal enters another cell. Each terminal has a unique identity by which it is identified by the mobile network and to which charges are assigned, Fig. 2, (2).

Terminal mobility is offered by the first generation of mobile networks, for example in the Nordic mobile telephone system, NMT, and the total access communication system, TACS. Terminal mobility offers mobility to users, but this mobility is limited to the coverage area of the mobile network, Fig. 3. The next step towards full mobility requires user identification.

Identification of users

Personal mobility involves identification of users without requiring any fixed relation between the terminal and the network, Fig. 2 (3). This means that a user may access telecommunication services on any terminal.

Furthermore, each user is given a service profile which specifies the services avail-

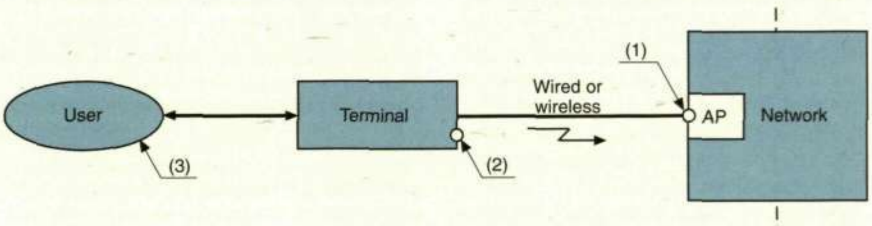


Fig. 2
Identification of lines, terminals and users in telecommunication networks

(1)	Line identity
(2)	Terminal identity
(3)	User identity
AP	Access port

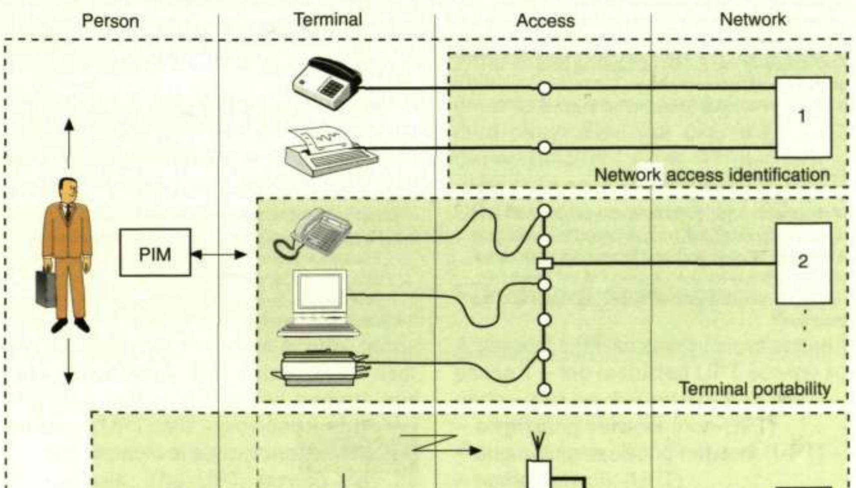


Fig. 3
Network access identification, terminal portability, terminal mobility and personal mobility

able, e. g. special treatment of incoming calls.

GSM (global system for mobile telecommunications) is a wireless system in which the identity of the user is given by a removable subscriber identity module, SIM, which can be inserted in any GSM terminal the user intends to use. GSM user mobility is thus a form of personal mobility restricted to the GSM environment, Fig. 3.

Definition of the UPT concept

Universal personal telecommunication, UPT, is a service that has been defined as follows in a draft by the ITU-TS SG 1 (see Box A):

"UPT enables access to telecommunication services while allowing personal mobility. It enables each UPT user to participate in a user-defined set of subscribed services and to initiate and receive calls on the basis of a personal, network-transparent UPT number

Box A

ITU-TS

Study Group 1

Study group 1, SG 1, (former CCITT SG I) handles questions related to services: service definitions, service operation, principles of service interworking, user quality and human factors. Services in SG 1 are to be defined and described from the user's point of view. SG 1 is thus the initiator and has the main responsibility for the UPT service within ITU-TS. SG1 has produced the first UPT standard: F.850 Principles of Universal Personal Telecommunication (UPT).

Study Group 2

Study group 2, SG 2, (former CCITT SG II) is responsible for questions related to network operation, including routing, numbering, network management and service quality. The services quality area includes traffic engineering, operational performance and service measurements. SG 2 has produced Draft Recommendation E.168, Application of the E.164 numbering plan for UPT and Draft Recommendation E.174, Routing for Universal Personal Telecommunication, UPT.

Study Group 3

Study group 3, SG 3, (former CCITT SG III), which is responsible for questions related to tariff and accounting principles, has proposed to investigate what kind of charging and accounting principles are appropriate for the provision of UPT services. Expected results are new recommendations in the D-series.

Study Group 7

Study group 7, SG 7, (former CCITT SG VII) handles questions related to data communication networks. In a longer perspective, the coverage of the UPT service will be extended to include public data networks.

Study Group 11

Study group 11, SG 11, (former CCITT SG XI) is responsible for questions related to switching and signalling. ITU-TS has been in a leading position in the intelligent network (IN) standardisation work. Together with SG 13, SG 11 has produced the first recommendations on IN in the Q.12xx series. The UPT service is to some extent included in the Capability Set 1, CS-1. A description of the UPT service in IN can be found in Draft Recommendation Q.1219. SG 11 has decided to plan, coordinate and standardise the signalling, call handling and management needed to support UPT in question 7, Signalling, call handling and management requirements for UPT, in the 1993-1996 study period. Task objectives include

- Stage 3 definition of UPT signalling for digital subscriber signalling system No. 1, DSS 1; for digital mobile access signalling, and for ISUP (ISDN services user part) enhancements for UPT
- Requirements for UPT functional distribution within fixed and/or mobile (land, satellite) networks and also with respect to multiple service providers.

Other UPT-related questions are

- Long-term intelligent network architecture. Future services – for example, multi-media utilising broadband-ISDN, B-ISDN, mobile communications and UPT service – will affect the long-term intelligent architecture. Also, associated emerging technologies such as IN-TMN (intelligent network – telecommunication management network) and IN-ODP (intelligent network – open distributed processing) have an affect on the long-term IN architecture
- Intelligent network capability sets. Seen from the UPT service's perspective, the study in this question will be focused on service creation for provision of the UPT service across multiple networks. The work on CS-2 and CS-3 (capability set 2 and 3) is expected to be completed in 1994 and 1996 respectively.

Study Group 12

Study group 12, SG 12, (former CCITT SG XII) handles questions related to end-to-end performance of networks and terminals. Wireless access is an important access method for the UPT service because it enhances personal mobility. SG 12 has therefore proposed to study a completely new question on the subject: Transmission performance of wireless personal communication systems.

Study Group 13

Study group 13, SG 13, (former CCITT SG XVIII) is responsible for studies related to general network aspects and the initial studies of the impact of new system concepts with far-reaching consequences. Together with SG 11, SG 13 has produced the first recommendations on IN in the Q.12xx series. In the 1993-1996 study period, SG 13 has proposed to study UPT performance. A Draft Recommendation, I.114, Vocabulary of terms for UPT, has been produced as well as Draft Recommendation I.373, Network capabilities to support universal personal telecommunication, UPT.

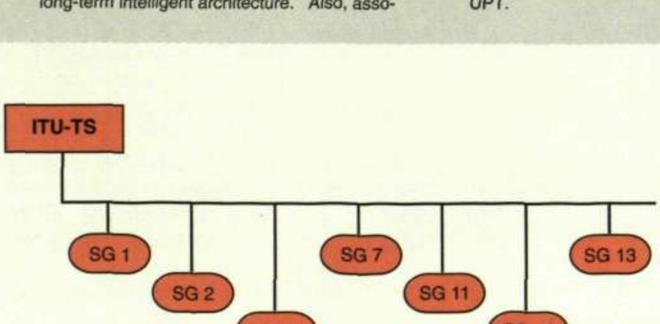


Fig. A
ITU-TS study groups (SGs) involved in the standardisation of UPT during the 1993-1996 study period

Fig. 4
A basic DTMF phone is all you need to inform the network of your service preferences and present whereabouts

across multiple networks on any fixed or mobile terminal, irrespective of geographical location, limited only by terminal and network capabilities and restrictions imposed by the network operator. Calls to UPT users may also be made by non-UPT users."

ETSI NA 7 (network aspects 7) has agreed on this definition, which has also been proposed by the ANSI (American National Standards Institute). It should be noted that the last sentence of the quotation applies to ETSI only.

This means that the fixed association between terminal or network access and user identification is removed. The identification of UPT users is treated separately from the addressing of terminals and network access points, in order to offer UPT users the capability to make and receive calls on any terminal and at any location, Fig. 3. The UPT user is personally associated with his/her own UPT number. This number is used as the basis for making and receiving calls, Fig. 2, (3). The UPT subscription is also charged on the basis of the user's UPT number.

The long-term goal of the UPT service is that a UPT user can use a multi-function integrated circuit (IC) card where the UPT subscription is one application. Because of the principle applied – that of identifying a person – the IC card is designated personal identity module, PIM.

UPT service description

UPT is intended for use in conjunction with other teleservices – telephony, facsimile and data, for example – to make the range of services offered both universal and personal.

A *universal teleservice* will allow access from multiple networks and provide terminal independence. In principle, a UPT user may use any terminal for making and receiving UPT calls – provided that the terminal is capable of supporting the requested services. The UPT service may be accessed by UPT users across multiple networks; for example, PSTN, ISDN, PLMN (public land mobile network) and PDN (public data network). Another dimension is operation over private telecommunication networks, PTNs, interworking with the public network.

A central part of the UPT service provision and use is the concept of *personal UPT service profiles* in which the service features and facilities subscribed to are specified. The UPT user can personalise, within some limits, his/her UPT service by modifying the content of the service profile.

A UPT number identifies each UPT user and is employed by the calling party to reach a specific UPT user. Multiple UPT user subscriptions are catered for in cases where a UPT user has more than one UPT number for different applications – for instance, a business UPT number and another UPT number for private calls. A company may have several UPT users – its employees – each with a personal UPT number and an associated UPT service profile.

Charging is associated with the UPT user's UPT number irrespective of the terminal or network used by the UPT user. This means that personal charging can be applied with one single bill covering all UPT services.

In order to facilitate and automate interaction between the UPT user and the UPT service, a UPT access device has been proposed, containing the UPT user identification. This device may be the size of a credit card.

Finally, a UPT user shall be guaranteed that his/her using the UPT service will involve no risk of violated privacy or fraudulent use of his/her identity. In principle, third party users (for example, terminal owners) shall not suffer with respect to privacy or freedom of action, as a result of UPT activities carried out by UPT users.

Functional architecture

A general UPT functional architecture for phase 1 – the restricted UPT service scenario – can be divided into four parts:
- originating network (non-UPT)
- originating/assisting network (UPT)
- home network (UPT)
- terminating network (non-UPT).

Fig. 5 shows the UPT functional architecture that is based on the intelligent network, IN, functional model CS-1 (capability set 1), defined by ITU-TS in Recommendation Q.1204. The IN func-

tional model can be divided into three logical groups
- call control related functions
- service control related functions
- management related functions.

The functional entities in the logical groups of the IN functional model (with respect to UPT) are briefly described below. For a general description of IN, see references.[13, 14, 26]

Call control related functions

The *call control agent function, CCAF*, provides access for both UPT and non-UPT users to the UPT service. This means that CCAF is the interface between the user and the network call control functions. The CCAF can be represented by a terminal.

A large amount of information has to be exchanged between the UPT user and the system; for example, information related to access, identification and authentication. Existing signalling systems do not support transfer of the required information, and that is why the new UPT-related information has to be transferred manually by in-band, dual tone multifrequency DTMF, signalling for UPT phase 1. From a practical point of view this is no limitation today, since DTMF phones are required for most additional telephone services, Fig. 4. Of course, a separate DTMF device can be used.

The *call control function, CCF*, which provides call/connection processing and control in the network,
- establishes, manipulates and releases call/connection instances according to information received from CCAF
- provides a trigger mechanism to access the IN functionality required for the UPT service, i.e. passing of events to the service switching function, SSF.

The signalling system between CCF and SSF has to transfer calling line identity to the SSF, and UPT user identity to the alerted terminal. In some cases, CCF may also be required to convert decadic pulses into DTMF signals. The management of CCF (e.g. trigger mechanisms) is handled by a service management function, SMF.

The *service switching function, SSF*, provides a set of functions needed for interaction between CCF and the service control function, SCF. SSF extends the logic of CCF, handles service control triggers and manages signalling between CCF and SCF. SSF also modifies the call/connection processing functions (in CCF), required to process requests under the control of SCF. The management of SCF is handled by an SMF.

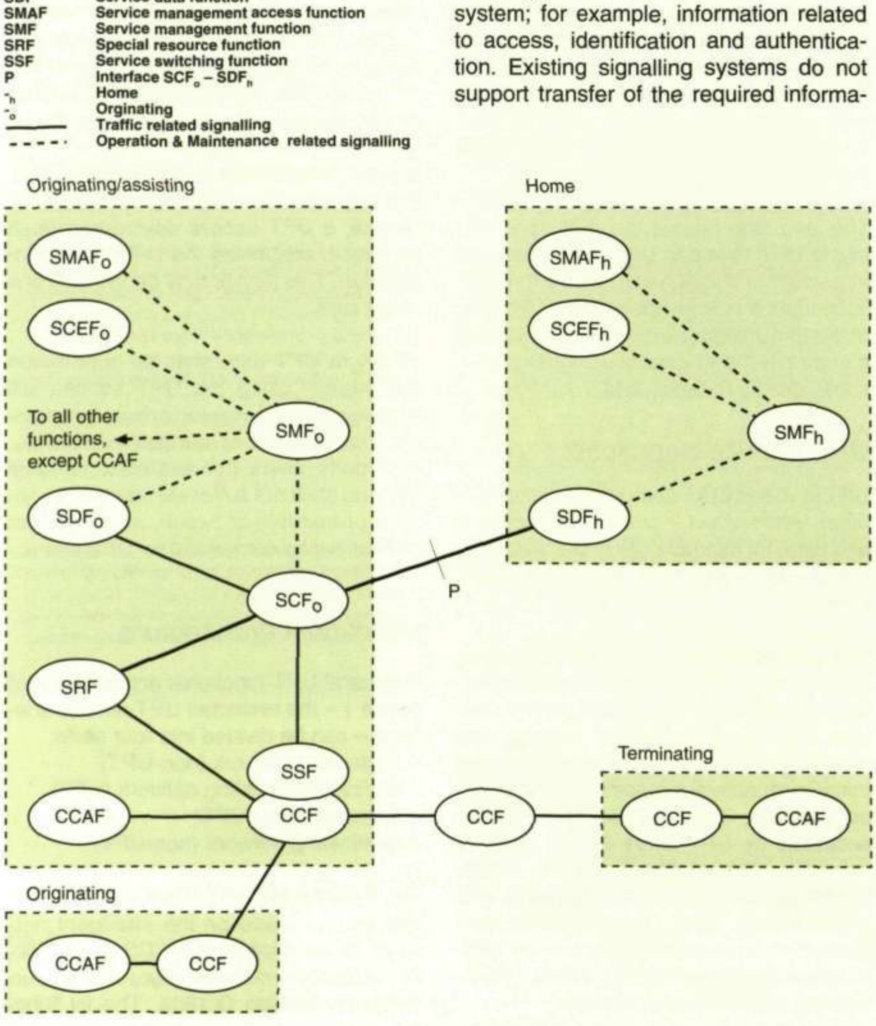
Service control related functions

The *service control function, SCF*, contains the logic and processing capability required to handle the UPT service. After a UPT call has been triggered in SSF, a request for instructions is sent from SSF to SCF. SCF performs UPT service control and responds to the request with instructions back to SSF.

SCF interfaces and interacts with SSF/CCF, the special resource function (SRF), and the service data function (SDF). SDF can be located in two different networks: in the originating/assisting network (SDF_o) and in the home network (SDF_h). SCF has real time access to SDFs in the execution of the UPT service. The

Fig. 5
Functional architecture of UPT according to ETSI phase 1 draft recommendations

CCAF	Call control agent function
CCF	Call control function
SCEF	Service creation environment function
SCF	Service control function
SDF	Service data function
SMAF	Service management access function
SMF	Service management function
SRF	Special resource function
SSF	Service switching function
P	Interface SCF_o – SDF_h
$-_h$	Home
$-_o$	Orginating
——	Traffic related signalling
- - - -	Operation & Maintenance related signalling



interface between SCF_o and SDF_h, which is indicated by P in Fig. 5, is needed to achieve appropriate information about the UPT user; for example, location information and service profile. The protocol between SCF_o and SDF_h is specified in CS-1 as the transaction capabilities application part, TCAP, of signalling system No. 7. Management of SCF is handled by an SMF.

The *service data function, SDF,* contains a considerable amount of UPT subscriber data and network data. SDF interfaces and interacts with SCFs. SDF_o stores

- a list of agreements, which indicates the identity of all the service providers whose subscribers are allowed to access the UPT service in SDF_o's network
- a list of service limitations resulting from agreements with service providers, or network limitations
- information related to the management of the UPT service.

SDF_h provides

- all data related to the UPT user; for example, location information, service profile and authentication information
- access control functionality to check whether requests received from remote entities are authorised or not
- authentication of the UPT user
- credit limit checks in real time for UPT calls.

Management of SDF_o and SDF_h is handled by an SMF_o and SMF_h respectively.

The *specialised resources function, SRF,* provides the specialised resources required for the execution of the UPT service. DTMF digit receivers, announcement machines and conference bridges are typical physical entities in SRFs. The SRF, which interfaces and interacts with SCF, SSF and CCF, is managed by an SMF.

Management related functions

The *service creation environment function, SCEF,* is the function where the UPT service is defined, developed, tested and input to SMF. Output from this function can be service logic, service management logic, service data and service trigger information.

The *service management access function, SMAF,* provides an interface between service managers and SMF. Service managers handle the UPT service in SMF through this interface.

The *service management function, SMF,* plays a central role for deployment and provision of the UPT service and for support of ongoing operation, coordinating different SCF and SDF instances by

- receiving billing and statistical information from the SCFs and making the information available to authorised service managers through SMAF
- modifying service data in SDFs
- managing service-related information in SRF, SSF and SCF.

Numbering scheme

In UPT, each user is identified by a unique UPT number to which some service aspects can be assigned. Firstly, the UPT number must be easily recognisable and distinguishable from ordinary (non-UPT) numbers. Secondly, the number should be as short as possible in order to minimise the number of digits the calling party must dial. Thirdly, and mandatory, it shall be

Fig. 6
The Ericsson service management system, SMAS, is an example of an efficient tool for creation and management of UPT services

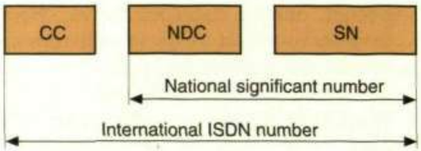


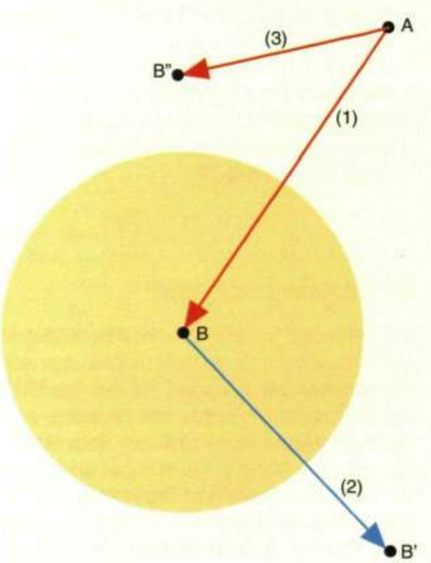
Fig. 7
Structure of ISDN number

CC	Country code
NDC	National destination code
SN	Subscriber number

Fig. 8
Principle of location-related charging in UPT

A	Calling party
B	Home location, called UPT user
B'	Registered location, called UPT user
B"	Registered location, called UPT user

Home location area
Charged to A-party
Charged to B-party



possible to dial a UPT number from any terminal.

The numbering scheme for UPT is currently being standardised in the ITU-TS Draft Recommendation E.168. The numbering structure in E.168 is based on the ITU-TS Recommendation E.164, which is the numbering plan for ISDN.

In Recommendation E.168, three different numbering schemes are proposed for UPT: a home-related scheme, a country-based scheme, and a global country-based scheme. In the following, only the country-based scheme is discussed.

In the *country-based scheme*, the E.164 number structure shown in Fig. 7 is – for UPT purposes – interpreted as

CC	Country code
NDC	UPT indicator/service provider indicator
SN	Subscriber number.

An international UPT number consists of CC + NDC + SN, while a national UPT number consists of NDC + SN. This is possible because the UPT information is placed in the NDC. The structuring of NDC is a national matter and may allow more than one UPT service provider in the same country.

Charging and billing

In UPT, charging and billing is associated with the UPT user's UPT number. Even if the UPT user accesses different networks from different terminals in different countries, preferably only one bill should be presented to the UPT subscriber by the UPT service provider.

The following types of charge may apply to UPT:
- subscription related charges, e. g. monthly fee
- subscription management related charges, e. g. for service profile modifications
- location related charges.

Location related charges
In PSTN today, as a rule only the calling party – the A-subscriber – is charged for the calls he/she makes. In GSM, call charging is split between the calling party and the called party, the B-subscriber. The calling party is charged for the part of the communication up to the called party's home location, while the called party is charged for the roaming part of the communication.

In UPT, location-related split charging applies. In other words, charging depends on the actual location of the calling and called parties, and the call charge may be split between them. Three different cases are discussed.

The calling party, A, (UPT or non-UPT) is charged for the part of the call from his/her own physical location to the home location area of the called UPT user, B. Thus, if the called UPT user is registered at his/her home location he or she will not be charged at all, Fig. 8, (1).

If the called UPT user is registered outside his/her home location, B', and the charge for the call to the registered location would be significantly higher compared with a call to B, then the call charge is split between the calling and called party:
- The calling party is charged for the call from his/her actual location to the called UPT user's home location, B, Fig. 8 (1)
- The called UPT user is charged for the roaming part of the call, from his/her home location, B, to the current registered location, B', Fig. 8, (2).

If the called UPT user is registered outside his/her home location, B", and the charge for the call would be significantly lower compared with a call to the home location, then the calling party, A, is charged for the whole connection, Fig. 8, (3). The called UPT user is not charged at all.

Service features

A service feature can be characterised as a specific aspect of a service that can be used in conjunction with other services or service features. It is either a core part of a service or an optional part offered as an enhancement to a service. Both ITU-TS SG 1 and ETSI NA7 have specified a set of UPT service features for provision in phase 1, divided into the following three groups:
- essential UPT features
- optional UPT features
- UPT supplementary services.

A feature that is regarded as fundamental to the UPT service concept is classified as essential. Optional UPT features are those features that are part of the basic operation of UPT, but not considered essential.

UPT supplementary services are enhancements to the essential and/or optional service features. Table 1 shows the UPT service feature classification according to ITU-TS and ETSI.

Authentication
UPT user identity authentication is a service feature by which the UPT service provider verifies that the identity of the UPT user is the one claimed. This is a very important feature and may therefore be used in each UPT service feature as a separate procedure.

UPT service provider authentication is a service feature which enables the UPT user to verify that the identity of the UPT service provider is the one claimed. This is important, especially in a multi-provider environment.

Personal mobility
Registration for incoming calls, *InCall registration*, is a service feature where the UPT user makes a registration, from the current terminal address, that all incoming calls to the UPT user in question should be presented to that terminal address. The registration is made as an update of the UPT user's current terminal address. The registration to a specific terminal address may have a certain period of validity specified. A new InCall registration from the same UPT user will cancel the one made previously. Also, an explicit de-registration of incoming calls is possible. More than one UPT user can be registered simultaneously to the same terminal address.

Registration for outgoing calls, *OutCall registration*, enables a UPT user, from the current terminal, to register outgoing calls made from that terminal address. After successful registration, and as long as it is valid, all outgoing UPT calls from the terminal address will be charged to the UPT subscriber. The UPT user may specify the period during which the registration is to be valid; otherwise it must be terminated through a deregistration procedure.

InCall and OutCall registration can be combined in a single procedure called *AllCall registration*. If the duration of the registration is not specified, an explicit de-registration is required.

Linked registration combines the InCall and OutCall registration and prevents the UPT user from overriding any part of the linked registration through an InCall or OutCall registration. The same UPT user must explicitly deregister a linked registration, or override it through another linked registration.

The service features described above can also be invoked from any terminal address. They are then complementarily named 'remote', e.g. remote InCall registration.

Table 1
Classification of phase 1 UPT service features

UPT service features	Classification ITU-TS	ETSI
Authentication		
• UPT user identity authentication	Essential	Essential
• UPT service provider authentication	Essential	-
Personal mobility		
• InCall registration	Essential	Essential
• OutCall registration	Optional	In a later phase
• AllCall registration	Optional	In a later phase
• Linked registration	Optional	In a later phase
• Remote InCall registration	Optional	Essential
• Remote OutCall registration	Optional	In a later phase
• Remote AllCall registration	Optional	In a later phase
• Remote linked registration	Optional	In a later phase
• Multiple terminal address registration	Optional	In a later phase
• Variable default InCall registration	Optional	-
Service profile management		
• UPT service profile interrogation	Optional	Essential
• UPT service profile modification	Optional	Essential
• Access to groups of UPT service profiles	Optional	Optional
• UPT specific indications	Optional	Essential
Call handling, calling party		
• Outgoing UPT call	Essential	Essential
• Call importance indication	Supplementary	-
• Calling party identification restriction	Supplementary	In a later phase
Call handling – called party		
• InCall delivery	Essential	-
• Variable routing	-	Optional
• Call forwarding unconditional	-	Supplementary
• Call forwarding on busy	Supplementary	Supplementary
• Call forwarding on no answer	Supplementary	Supplementary
• Call forwarding on not reachable	-	Supplementary
• Call pick-up	Optional	In a later phase
• Calling party identification presentation	Supplementary	In a later phase
• Called party specified secure answering of incoming UPT calls	Optional	In a later phase
• Intended recipient identity presentation	Optional	In a later phase
Follow-on		
• OutCall follow-on	Optional	Essential
• Global follow-on	Optional	Essential
Help desk		
• Operator-assisted services	Essential	Optional

Fig. 9
The UPT service feature *variable default Incall registration* provides the network with information of your regular habits, making it possible to route all incoming calls to the terminal where you are most likely to be reached

Variable default InCall registration enables the UPT user to set up a default registration matrix of terminal addresses for incoming UPT calls, so that incoming UPT calls can be routed and handled differently according to time of day, day of week, calling party's identity, type of service, the number dialled, and according to 'on no answer' or 'on busy' conditions, as appropriate. A UPT user can modify his/her own matrix so that it fits, for instance, his/her regular travel routine or time schedule, Fig. 9.

Service profile management

For service management, the following service features are available to the UPT user:

- *UPT service profile interrogation* enables the UPT user to interrogate the current status of his/her own service profile. For example, location information or availability of services may be interrogated.
- *UPT service profile modification* allows the UPT user to modify the content of his/her own service profile. For example, password or default parameter values in the service profile can be changed.
- *Access to groups of service profiles* permits a UPT subscriber or an authorised representative of a group of UPT users to access, create, modify and interrogate their service profiles through service profile management procedures.
- *UPT-specific indications* provide a set of user-friendly standard announcements or indications for the UPT user. UPT-specific indications may be used as announcements in specific charging arrangements.

Call handling – Calling party

Outgoing UPT call allows a UPT user to make a single outgoing UPT call from any terminal and be charged for the call. This feature requires authentication for each UPT call attempt, unless a follow-on service feature is invoked, Fig. 10.

Calling party identification restriction, CPIR, prevents the calling party identity from being presented on the alerted terminal. Thus, the calling party stays anonymous to the receiver of the call.

Call handling – Called party

InCall delivery presents all incoming calls to the UPT user at the terminal address previously registered by InCall registration.

Call forwarding on no answer forwards incoming UPT calls to another line number or terminal address in case of a no answer condition at the registered termination point.

Calling party Identification presentation, CPIP, presents the calling party's identity on the alerted terminal. The calling party's identity may be the UPT number or name, rather than a terminal address.

Follow-on

Follow-on means that a UPT user can establish a session consisting of several service features without user authentication in each case.

OutCall follow-on permits a UPT user to make a sequence of outgoing UPT calls without making authentication procedures between the calls. When terminating a call, the UPT user indicates that another outgoing UPT call follows. This feature is limited to outgoing UPT calls.

Global follow-on enables the UPT user to make a sequence of service features without repeating the authentication procedure. Before completed disconnect, the user may indicate that other service features will follow. Compared with OutCall follow-on, this service feature is general and can be applied to any service feature.

Help desk

Help desk or *operator-assistance* is a standardised and comfortable way for a UPT user to contact a UPT service centre and be assisted; for example, when automatic UPT procedures are unavailable.

Future aspects of the UPT service

In the future, a large number of enhanced service features will be included in the UPT service. Some of these features are discussed below.

Service personalisation allows the UPT user to tailor the UPT service; for instance, to specify the preferred language for all dialogues independent of geographical location. This information is stored in the UPT user's service profile.

Advanced addressing capability (ETSI's designation is personal addressing) enables the UPT user to use his/her name, address or profession – or a combination of these items – instead of a number, to initiate an outgoing UPT call.

Ongoing call redirection enables the UPT user to put on hold an incoming or outgoing call in progress, and then take over the call at another terminal.

Calling party specified secure answering of calls to UPT users is a service feature by which the calling party (UPT user or non-UPT user) can specify that the answering UPT user must first successfully authenticate himself/herself before the call can be answered.

Call importance indication enables the calling party (UPT or non-UPT) to indicate to the called UPT user that the call is important.

Priority screening of incoming UPT calls permits the UPT user to specify how incoming UPT calls with priority are to be treated. Important calls can be directed to the UPT user's office telephone, while other calls can be directed to a secretary, Fig. 11.

Fig. 10
When you are abroad, a mobile phone puts you in contact with your home location. After authentication, all your personalised services are immediately available. Charges will be on your regular account

Security

The freedom given to UPT users to move from one terminal to another and access any other terminal from different networks also implies that attempts to fraudulently use their subscription can be made from any terminal in any network. It is therefore necessary that the UPT service provides sufficient security mechanisms. These security mechanisms should appear in a user-friendly way, which makes security features an extensive part of UPT.

Authentication is a process by which the claimed identity of an entity is verified by another entity. In UPT, authentication of the user to the service provider is required. This authentication can be either weak or strong.

A weak authentication mechanism is based on permanent data; for instance, a personal identification number, PIN, which is entered by the UPT user each time an authentication process is performed. The PIN length may be from 4 to 8 digits. A weak authentication mechanism is regarded as insufficient in UPT phase 1.

In strong authentication, authentication data is changed unpredictably for each authentication attempt and thus counteracts repeated attacks. Furthermore, access to the network is established through a UPT access device. The whole authentication process is divided into two parts:
- authentication of the user to the UPT device
- authentication of the UPT device to the system (network/service provider).

Authentication of the user to the device is made by the UPT user entering a local PIN, LPIN, to the UPT device. A valid authentication unlocks the UPT device. The LPIN has a length of, typically, 4 to 8 digits.

Authentication to the system is made by the UPT device. The UPT user initiates the procedures on the UPT access device while the device is connected to the microphone of a telephone. The UPT access device contains all individual and security-related algorithms and data. In-band DTMF signalling is used from the UPT device to a DTMF digit receiver in the special resources function, SRF. A security module in the home SDF_h contains the algorithms and keys for performing the procedures. After successful authentication, the UPT user is requested to proceed and select the desired service feature.

Fig. 11
In the future, *Priority screening of incoming calls* **will enable the UPT user to accept only important calls – when at lunch, for example. All other calls can be directed to an assistent or to an answering service**

Man-machine interface and UPT procedures

The man-machine interface (MMI) is an interface between the UPT user and the terminal or the UPT access device. Customer control of the UPT service applies. In phase 1 (restricted UPT service scenario) MMI and UPT procedures are limited to interactive voice prompting with DTMF signalling. DTMF signalling can be established from an ordinary DTMF push-button telephone set or from a separate DTMF device. In a later phase, other methods may be used: modems, different types of cards (magnetic strip cards and IC cards) and card readers. Fig. 12 shows an example of how MMI and UPT procedures may be implemented.

Access procedure
The UPT user performs an access procedure to obtain access to the UPT service.

Identification procedure
A UPT user performs an identification procedure to identify himself/herself to the UPT service provider. The UPT user enters some information, e.g. UPT number.

Authentication procedure
The authentication procedure is used by the UPT service provider to ensure that the calling or called party is the UPT user claimed. This is an important procedure and may therefore be used in conjunction with most of the other UPT procedures. Fig. 12 shows an example of the use of a weak authentication procedure.

Personal mobility procedures
Personal mobility procedures are related to the personal mobility of the UPT user and performed in order to ensure that the user can receive and make UPT calls. However, procedures for personal mobility do not include the making or receiving of UPT calls. Personal mobility procedures are provided for by InCall registration, OutCall registration, AllCall registration and linked registration.

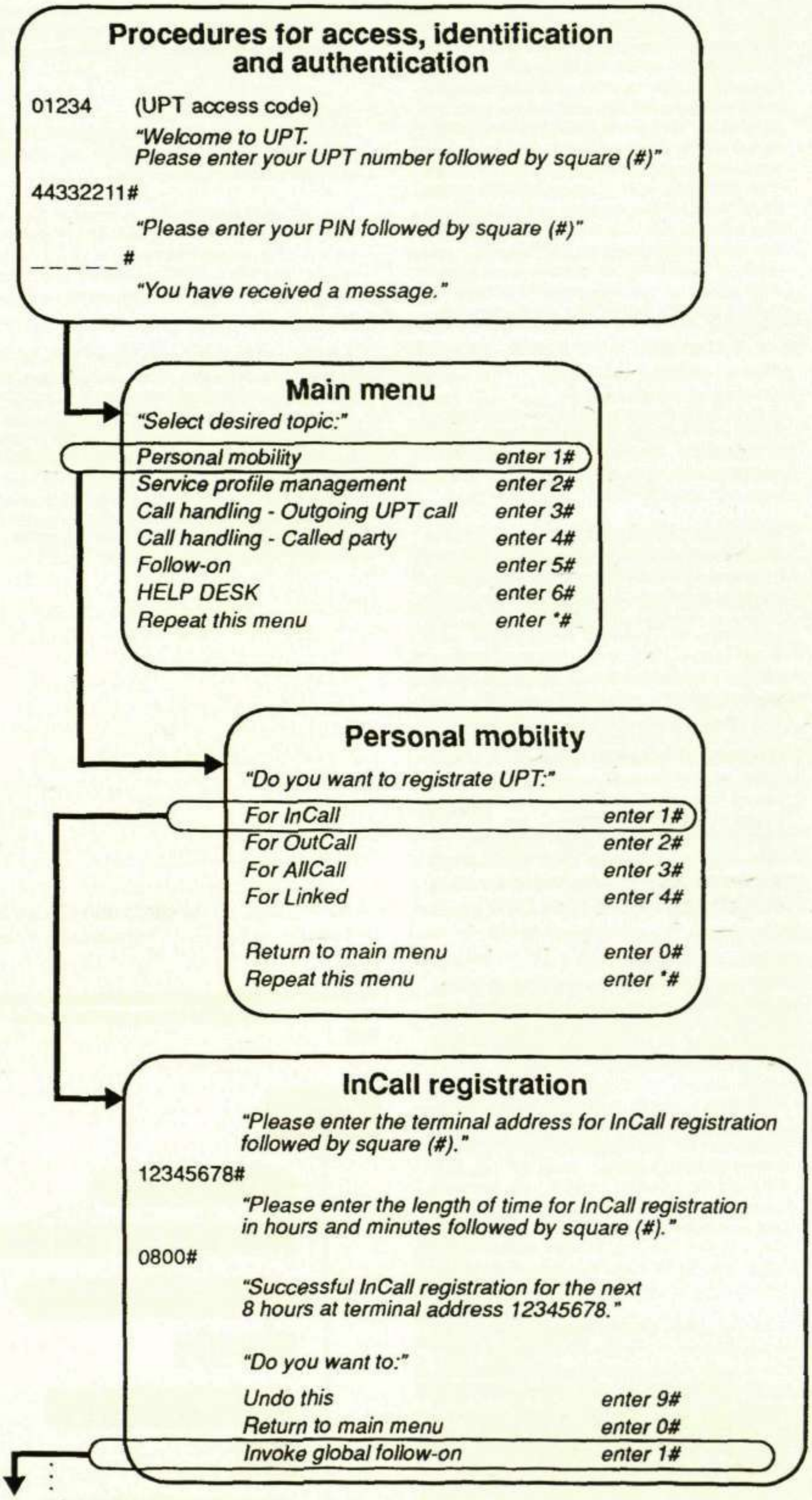


Fig. 12
Example of how MMI and previously described UPT procedures may be implemented. Different ways of presentation may be involved, depending on the access type. For instance, from an ordinary telephone set (telephony) the information is gathered through voice prompting. This means that the system plays announcements to the UPT user, who has to respond with DTMF signalling. Another access type, with a display, enables the UPT user to receive graphical information on the screen and respond by touching the screen, using a mouse or pressing buttons on a separate keyboard

Standardisation activities of UPT

The UPT service is subject to standardisation in national, continental and international standardisation bodies. Standardisation activities are primarily being carried out by the International Telecommunication Union – Telecommunication Standardisation Sector, ITU-TS, and the European Telecommunications Standards Institute, ETSI.

ITU-TS

ITU-TS is working on standardisation of UPT and UPT-related questions in a number of Study Groups (SGs), Box A. Fig. A shows the SGs involved during the 1993-1996 study period.

Study group 1

Study group 1, SG 1, (former CCITT SG I) handles questions related to services. It is responsible for service definitions, service operation, principles of service interworking, user quality and human factors. It is to be noted that services in SG 1 are to be defined and described from the user's point of view. This means that SG 1 is the initiator and has the main responsibility for the UPT service within ITU-TS.

The former CCITT SG I started a UPT study under question 35 in the 1989-1992 study period. SG 1 continues the work under question 7, Universal Personal Telecommunication (UPT), in the 1993-1996 study period. The following issues are proposed for consideration in question 7:

- user perception of the UPT service
- user interface
- interaction with fixed location and mobile service provision
- security of access to services
- privacy, especially of location and service data
- compatibility with existing and evolving network services
- international availability
- operational aspects of charging and recording of call data
- which features should be essential and which additional
- interaction with network-based supplementary services
- how to introduce UPT in a phased manner.

Box B

ETSI

The technical committee "Network Aspects", NA, is responsible for defining general aspects of all existing and new networks. This includes network modelling, architecture, definition of network functions and fundamental structure of user-network interfaces. NA is further divided into six technical sub-committees, five of which are involved in the standardisation of UPT.

Sub technical committee NA1
NA1 is in charge of user interfaces, services and charging. NA1 is responsible for producing an ETSI standard for UPT, entitled "Definition of a UPT phase 1 service".

Sub technical committee NA2
Questions related to numbering, addressing, routing and interworking are handled by NA2. NA2 is carrying out studies of UPT numbering and routing.

Sub technical committee NA4
NA4 is responsible for standardisation work on network architecture operations, maintenance, principles and performance. A large and important area for NA4 is management of the UPT service and related security matters which may impact both TMN and IN.

Sub technical committee NA6
NA6 is responsible for standardisation of intelligent networks. NA6 has not published any general IN standards but influences standardisation work on both CS-1 and CS-2. Based on the ITU-TS IN de facto standards, the NA6 also produces technical reports on specific implementational questions. A special group of experts within NA6 has been established to investigate security mechanisms for node-node communication.

Sub technical committee NA7
NA7, which is working more or less on a project basis, is responsible for:

- definition of a target service for UPT agreed in ETSI and harmonised with other bodies. Services are to be defined both from the user's and the operator's points of view
- identification of open issues and needs for standardisation
- guidelines and requirements for work on UPT in other ETSI groups
- development of a programme of coordinated actions to be performed towards the target UPT service (framework for a set of standards).

Technical committee SPS
Aspects of signalling, protocols and switching are studied in the technical committee signalling, protocols and switching, SPS. SPS is responsible for defining information flows, and the call handling sequence and signalling in the public network, including techniques for transfer of user-to-user information, user-to-node communication and inter-node communication.

The technical sub-committee SPS2 is responsible for signalling network and mobility applications. SPS2 is working on an ETSI standard "Application protocol for UPT", which is a stage 3 description of UPT phase 1.

Another technical sub-committee within SPS is SPS3, which is responsible for digital switching and produces what is called a European "core" intelligent network application protocol (INAP) by reducing options from the ITU-TS recommendation Q.1218. The "core" INAP, which will support the UPT service, has to be extended with respect to SCF – SDF operations.

Technical committee BT
The technical committee business telecommunications, BT, is responsible for studying business telecommunication aspects of private networks, public networks and their attached terminals. Private networking aspects within ETSI are handled by sub technical committee BT1. BT1 is studying the possibility of providing access to UPT via private telecommunication networks, PTNs. Access to UPT from PTNs is regarded as a security problem, because a private branch exchange, PBX, can make recordings of any DTMF signalling. That is why a weak UPT user authentication mechanism (based on a permanent personal identification number, PIN, and DTMF signalling) is not recommended.

Technical committee RES
Questions concerning radio communications equipment and systems (except those specifically allocated to other technical committees) are handled by the technical committee radio equipment and systems, RES.

Sub technical committee RES3 is responsible for the standardisation of digital European cordless telephone, DECT. DECT is a cordless telecommunication system which may provide radio access to the PSTN, ISDN, PLMN (public land mobile network), and PBX for high traffic density areas, including both voice and data services. RES3 is investigating the use of DECT terminals together with UPT. In the first phase of UPT, there is no new requirement for DECT, because all UPT procedures are supposed to be carried out with in-band DTMF signalling. However, in a later phase, when IC cards may be included in the UPT service – e.g. personal identification module, PIM – there is a need for coordination between UPT and DECT. The influence of UPT on the DECT IC card, which is called DAM (DECT access module), and the DECT signalling channel are example of new questions for RES3.

Technical committee SMG
Technical committee Special Mobile Group, SMG, is responsible for defining all aspects of digital cellular telecommunications systems including services/facilities, radio interface, network aspects and telematic services.

Sub technical committee SMG1 is responsible for services and facilities in GSM. Clearly, standardisation of UPT and standardisation of GSM are interdependent. On the one hand, UPT phase 2 will place requirements on GSM. On the other hand, GSM – in terms of standardisation and implementation – is ahead of UPT. It is therefore essential that future GSM work carried out in SMG1 will take UPT requirements into consideration.

SMG5 is a sub technical committee of SMG which is responsible for the standardisation of UMTS (universal mobile telecommunication system) i.e. the third generation of mobile telecommunication systems. UMTS will probably be the first telecommunication system in which the UPT service concept has been taken into consideration from the very beginning.

Technical committee HF
The technical committee human factors (HF) is responsible for user interface – i.e. MMI (man-machine interface) – standards for telecommunications equipment and services.

The sub technical committee HF1 is in charge of human factor aspects of teleservices and supplementary services including the preparation of human factor standards, recommendations and guidelines. The MMI is important to the entire UPT service provision and use. For instance, if users' perception of the UPT service is not positive it may have adverse effects on UPT service deployment.

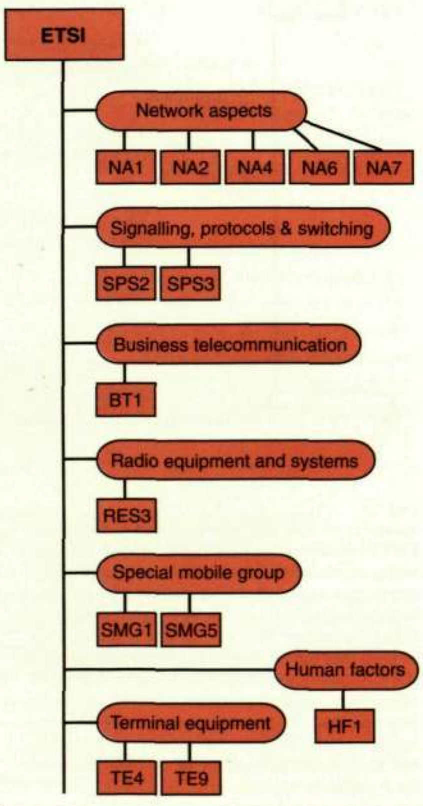
HF1 is responsible for the development of UPT procedures and guidelines for UPT announcements.

Technical committee TE
Technical committee terminal equipment, TE, is responsible for the standardisation of terminal equipment and terminal systems.

Sub technical committee TE4 handles the area of audiovisual and voice terminals. TE4 is responsible for the standardisation of a UPT access device, which is a DTMF sender for acoustical coupling to the microphone in the telephone handset.

Another technical sub-committee, TE9, is in charge of IC cards and the card-related parts of terminals used in telecommunications. For UPT it is proposed to use the concept of the multi-application card under development in TE9. In the future, the multi-application card may be a personal identity module, PIM, in which the UPT service is one application among others. TE9 has also proposed an authentication algorithm for the UPT device and system.

Fig. B
Technical committees and sub technical committees in ETSI involved in the standardisation of UPT



SG 1 has produced the first UPT Recommendation F.850, Principles of Universal Personal Telecommunication (UPT). F.851, Universal Personal Telecommunication (UPT) – Service Description (draft), which SG 1 produces in liaison with other study groups, is a stage 1 description of the UPT service.

Many ITU-TS study groups are involved in the standardisation of the UPT service. As a result of the work so far, the following major UPT recommendations (drafts or approved versions) exist (originating SG within brackets):
- F.850, Principles of Universal Personal Telecommunication (UPT) (SG 1)
- F.851, Universal Personal Telecommunication (UPT) – Service Description (SG 1)
- E.168, Application of the E.164 numbering plan for UPT (SG 2)
- E.174, Routing for Universal Personal Telecommunication (UPT) (SG 2)
- Q.1219, Intelligent Network user's guide to CS-1. (SG 11)
- Q.76, Service procedures for universal personal telecommunication functional modelling and information flows (SG 11)
- I.373, Network capabilities to support Universal Personal Telecommunication (UPT) (SG 13)
- I.114, Vocabulary of terms for UPT (SG 13)

ETSI

The European Telecommunications Standards Institute (ETSI) works with standardisation of UPT in a number of technical committees and sub-committees, Box B. Fig. B shows the structure of ETSI from a UPT point of view. ETSI quite early recognised the need to start standardisation of UPT, and a new technical sub-committee called NA7 (working more or less on a project basis) was established in 1990.

Other standardisation bodies

ITU-RS

Standardisation work on the future public land mobile telecommunication systems, FPLMTS, was started within the CCIR (International Radiocommunication Consultative Committee) around 1985 and is now being continued within the International Telecommunication Union – Radiocommunication Sector, ITU-RS. Like UMTS (universal mobile telecommunication system), FPLMTS is a third-generation mobile system, scheduled to start service around the year 2000. In its draft of a new recommendation, Study Group 8, SG 8, (former CCIR Task Group 8/1, TG 8/1) has stated that FPLMTS will support UPT since it is standardised by ITU-TS.

ANSI

The American National Standards Institute, ANSI, technical sub-committee T1P1, is working on a new concept called personal communications which includes personal communications service, PCS, and personal communications system. The key attributes of PCS are
- personal mobility via UPT service
- terminal mobility via wireless access mobility services
- ubiquitous coverage and connectivity to the public network
- high quality and availability
- wired and wireless access
- utilisation of IN capabilities.

Evolution and service standardisation of UPT

UPT is expected to materialise during the next ten-year period and will follow an evolutionary path that is strongly influenced by market needs and technological advances. Both ETSI and ITU-TS have therefore proposed a phased approach to the standardisation of UPT. Provision of the UPT service will first start with a simplified set of service features. More advanced scenarios will be introduced later.

In accordance with ETSI notation, the standardisation of the UPT service is split into the following phases:
- Phase 1, restricted UPT service scenario
- Phase 2, basic UPT service scenario
- Phase 3, enhanced UPT service scenario.

Fig. 13 shows the phased approach to the UPT service standardisation. However, it should be borne in mind that future technological and market developments may give rise to new, as yet unforeseeable, evolutionary phases of the UPT service. In these circumstances, the standardisation work has to be rearranged to correspond with the UPT implementation phases.

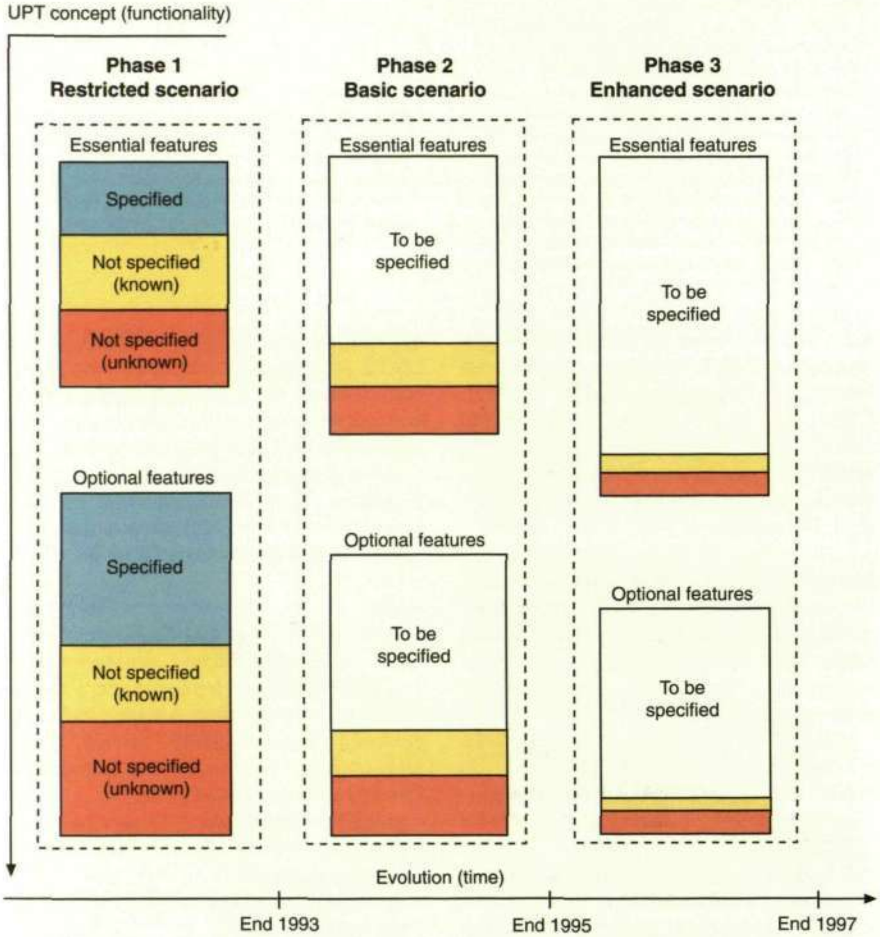


Fig. 13
Evolution and service standardisation phases of UPT in ETSI and ITU-TS. The time schedule is according to the present ETSI work programme

A short description of each phase of the UPT service evolution with its corresponding time frame is presented in the following.

Phase 1 – restricted UPT service scenario
The first phase, phase 1, of the UPT service is a restricted short-term scenario. It has restrictions in networks, service features, security and user friendliness. The restricted UPT service scenario consists of a set of service features that can be implemented in existing networks without any major changes to these networks. Phase 1 is therefore basically restricted to PSTN, ISDN and, possibly, PLMN.

Only telephony service (voice) is provided for. Furthermore, only a subset of UPT service features is specified in this phase. The specified features are classified either as essential or optional, depending on whether the feature in question is necessary to the UPT service provision or not. The target date for phase 1 – restricted UPT service scenario – standards is at the end of 1993, according to the ETSI work programme, Fig. 13.

Phase 2 – basic UPT service scenario
Phase 2, which is called the basic UPT service scenario, will incorporate more service features and networks, moving towards full universal service availability, terminal independence and operation across multiple networks. IN technology may be used to implement this scenario. The basic UPT service scenario also provides a set of data services.

According to ETSI, the UPT service will even be provided in connection-oriented data networks and in the GSM network. Furthermore, IC cards and card readers are the subject of standardisation. The target date for phase 2 – basic UPT service scenario – standards is the end of 1995, Fig. 13.

Phase 3 – enhanced UPT service scenario
Phase 3, an enhanced UPT service scenario, is the long-term scenario of the UPT service. In the future it is likely that several technological and market developments may give rise to evolutionary phases of the UPT service, which cannot be foreseen today. The target date for phase 3 – enhanced UPT service scenario – standards is the end of 1997, Fig. 13.

Summary

Universal personal telecommunication, UPT, is a new service concept, which has relatively recently been studied and developed. In UPT, the fixed association between terminal or network access and user identification is removed. UPT supports personal mobility and allows the UPT user, who is associated with a personal and unique UPT number, to make and receive calls on any terminal and at any location.

UPT will follow an evolutionary path, strongly influenced by market needs and technological advances. It is generally expected to be one of the most important services in the future.

References

1 CCIR Draft New Recommendation [DOC. 8/51]. *Framework for Services Supported on Future Land Mobile Telecommunication Systems (FPLMTS).* The International Radiocommunication Consultative Committee, 1992 (January 30).

2 CCIR Draft New Recommendation [DOC. 8/52]. *Future Land Mobile Telecommunication Systems Network Architecture.* The International Radiocommunication Consultative Committee, 1992 (January 30).

3 CCITT Draft Recommendation E.168. *Application of E.164 Numbering Plan for UPT.* International Telegraph and Telephone Consultative Committee, 1992.

4 CCITT Draft Recommendation E.174. *Application of E.174 Routing for Universal Personal Telecommunications (UPT).* International Telegraph and Telephone Consultative Committee, 1992.

5 CCITT Draft Recommendation F.850. *Principles of Universal Personal Telecommunication (UPT).* The International Telegraph and Telephone Consultative Committee (Version 4) 1992.

6 ITU-TS Draft Recommendation F.851. *Universal Personal Telecommunication (UPT) - Service Description.* International Telecommunication Union-Telecommunication Standardization Sector, (Version 9) 1993 (April 30).

7 CCITT Draft Recommendation I.373. *Network Capabilities to support Universal Personal Telecommunication (UPT).* The International Telegraph and Telephone Consultative Committee, 1992.

8 CCITT Draft Recommendation Q.1204. *Intelligent Network Distributed Functional Plane Architecture.* The International Telegraph and Telephone Consultative Committee, 1992 (March 17).

9 CCITT Draft Recommendation Q.1211. *Introduction to Intelligent Network Capability Set - 1.* The International Telegraph and Telephone Consultative Committee, 1992 (March 17).

10 CCITT Draft Recommendation Q.1215. *Physical Plane for Intelligent Network CS-1.* The International Telegraph and Telephone Consultative Committee, 1992 (March 17).

11 CCITT Draft Recommendation Q.1219. *Intelligent Network User's Guide for CS-1.* The International Telegraph and Telephone Consultative Committee, 1992 (March 17).

12 CCITT Recommendation E.164. *Numbering Plan for the ISDN Era.* The International Telegraph and Telephone Consultative Committee, 1988.

13 van Hal, P., van der Meer, J. and Salah, N.: *The Service Script Interpreter, an Advanced Intelligent Network Platform.* Ericsson Review *67*(1990):1, pp. 12-22.

14 Ljungblom, F.: *A Service Management System for the Intelligent Network.* Ericsson Review *67*(1990):1, pp. 32-41.

15 Network Aspects (NA): *Universal Personal Telecommunication (UPT) Specification of the Security Architecture for UPT Phase 1.* Valbonne Cedex 1993, European Telecommunications Standards Institute, Draft ETSI Technical Report DE NA-71401 (Version 0.2.1).

16 Network Aspects (NA): *Universal Personal Telecommunication (UPT) Requirements on Information Flows and Protocols.* Valbonne Cedex 1992, European Telecommunications Standards Institute, Draft ETSI Technical Report TC-TR NA-70301 (Version 0.1.0).

17 Network Aspects (NA): *Universal Personal Telecommunication (UPT): General Service Description.* Valbonne Cedex 1992, European Telecommunications Standards Institute, Draft ETSI Technical Report ETR NA-70201 (Version 1.2.0).

18 Network Aspects (NA): *Universal Personal Telecommunication (UPT): Phase 1 (Restricted UPT Service Scenario) Service Aspects: Guidelines.* 1992, European Telecommunications Standards Institute, Draft ETSI Technical Report TC-TR NA-71101 (Version 0.6.0).

19 Network Aspects (NA); *Universal Personal Telecommunication (UPT) Phase 1(Restricted UPT Service Scenario) Service Aspects of Charging, Billing and Accounting.* Valbonne Cedex 1992, European Telecommunications Standards Institute, Draft ETSI Technical Report TC-TR NA-71202 (Version 0.3.0).

20 Network Aspects (NA); *Universal Personal Telecommunication (UPT): Architecture and Functionalities for Interworking.* Valbonne Cedex 1993, European Telecommunications Standards Institute, Draft ETSI Technical Report DTR NA-70303 (Version 0.3.0).

21 Network Aspects (NA); *Universal Personal Telecommunication UPT Phase 1: Architecture and Functionalities for Interworking.* Valbonne Cedex 1992, European Telecommunications Standards Institute, Draft ETSI Technical Report TC-TR NA-71303 (Version 0.4.0).

22 *Questions Allocated to Study Group 1 - (Service Definition) - for the 1993-1996 Study Period.* Geneva 1993, International Telecommunication Union-Telecommunication Standardization Sector, COM 1-1-E.

23 *Report of the 8th Meeting of STC NA7 Universal Personal Telecommunication, Held in Madrid, Spain 23-27 November 1992.* Valbonne Cedex 1992, European Telecommunications Standards Institute, ETSI STC NA7(92)19.

24 *Report of the 9th Meeting of STC NA7, Universal Personal Telecommunication, Held in Vienna, Austria 22-26 March 1993.* Valbonne Cedex 1993, European Telecommunications Standards Institute, ETSI STC NA7(93)21.

25 Sundborg, J.: *A Conceptual Model of Universal Personal Telecommunication Service Provision and Use.* Master's Thesis. Helsinki University of Technology, Faculty of Electrical Engineering. Helsinki 1993.

26 Söderberg, L.: *Architecture for Intelligent Networks.* Ericsson Review *66*(1989):1, pp. 13-22. p. 13-22.

27 Söderberg, L.: *Evolving an Intelligent Architecture for Personal Telecommunication.* Ericsson Review 70(1993):4 pp. 156-171.

28 Technical Sub-Committee T1P1: *System Engineering Working Document for Personal Communications.* USA 1992, American National Standards Institute, T1P1.2/92-002R3, T1P1.3/92-002R3 (Draft 6).

29 *Terms of Reference for Technical Committees & Sub-Committees of the European Telecommunications Standards Institute.* Valbonne 1992, European Telecommunications Standards Institute, TOR 10/92.

30 *Universal Personal Telecommunication (UPT) Phase 1, UPT Supplementary Services.* Valbonne Cedex 1993, European Telecommunications Standards Institute, Draft ETSI Technical Report DTR NA-71210 (Version 1.0.0).

31 *Universal Personal Telecommunication (UPT): Supplementary Services (release 2).* Valbonne Cedex 1993, European Telecommunications Standards Institute, Draft ETSI Technical Report DTR NA-70210 (Version 1.0.0).

32 Vandenameele, J., Thieffry, J.-B. and Decuypere, H.: *UPT, A New Dimension in Telecommunications Provided by IN.* 1992 International Zurich Seminar on Digital Communications – Intelligent Network and their Applications, ETH Zurich, Switzerland, March 16-19, 1992. Berne, Switzerland 1992, IEEE Switzerland Chapter on Digital Communications. pp. 41-54.

Evolving an Intelligent Architecture for Personal Telecommunication

Lennart Söderberg

Personal telecommunication in various forms promises to give users a new quality in telecommunication and is therefore attracting the market's keen interest. Terminal manufacturers see a new mass market for pocket telephones, cellular operators see a new niche to expand their markets, and traditional operators see new opportunities to increase revenue from their networks. Some see a role for new operators to provide cheap and flexible accesses. Others see a role for new service providers to develop new service offerings independently of the network operators. Legislation, market forces and the progress of standardisation will determine the outcome.
The author gives an overview of requirements from the various players on the market, presents a vision of a future freedom of personal telecommunication and outlines a matching architecture that integrates intelligent network and cellular technologies. The proposed architecture provides for stepwise implementation of personal telecommunication, but also permits compromises during the introductory phases.

The 1990s is a dynamic time for the telecommunications industry. Regulatory, technical and commercial issues are often in conflict, causing established practices to be questioned and inciting new approaches. Personal telecommunication, PT, is such an issue; it promises new revenues from existing network investments, and opportunities to reduce costs by introducing new technologies in the network.

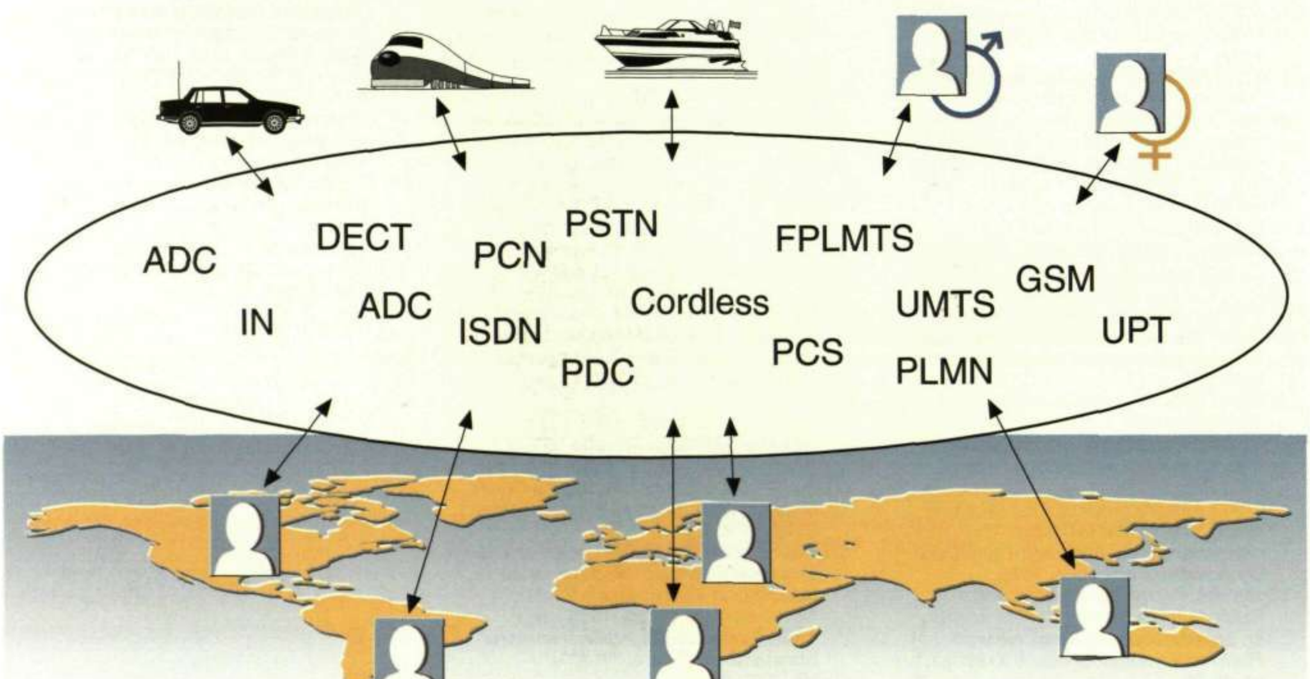
Personal telecommunication means getting people on speaking terms. You request the network to connect you to the person you want to reach – not to a place or a terminal. You need not know where this person is for the moment – the network will find out.

The objective of personal telecommunication is ease of use, choice of price and means of communicating with anyone, anytime, anywhere – whether at work, at home or on the move.

When the ordinary telephone number can give this freedom, the telephone network will have become a service-independent network for personal telecommunication – a new mode of operation – combining the best from different telecom technologies: mobility from cellular networks, connectivity from the global telephone network, and flexibility from the Intelligent Network (IN) concept.

ETSI and ITU are working on two concepts for the longer term, beyond 2000-2005: the universal mobile telecommunications system, UMTS, and future public land mobile telecommunication systems, FPLMTS. Apart from the cellular stan-

Fig. 1
The term personal telecommunication expresses the need for people to communicate with anybody, anywhere, independent of access. The picture illustrates that there are many different standards to be considered when defining the service



LENNART SÖDERBERG
Ericsson Telecom AB

dards, there are two international short-term approaches to personal telecommunication. One is universal personal telecommunication, UPT, for which ETSI/ ITU-TS is making an concentrated effort to release standards stepwise over the period from 1993 to 1997. Service definitions are related to a personal number, the UPT number, independent of terminal or network port numbers. The primary objective is to enhance current telephone networks with UPT capabilities based on IN standards. The second approach, personal communication services, PCS, is a massive effort in the US to introduce low-cost, high-volume, pocket telephones, envisaging a market of 50 million users in the next ten years. Special licences are being allocated and field trials are already under way. Bellcore is releasing their AIN Rel 0.2 standard in 1993. ANSI is working on the ITU-TS standards CS1 and CS2. Radio access standards are developed towards microcells and picocells.

The second generation of macro cellular network standards – the European global system for mobile communication, GSM, the American digital cellular standards, here called ADC, and the Japanese personal digital cellular standard, PDC – offers radio access in support of high-speed mobility for public land mobile networks, PLMN. A network conforming to any of these standards will offer personal telecommunication with identical network behaviour (feature transparency). However, to turn these approaches into successful PT mass-market solutions, prices have to go down, the acceptable terminal density has to go up (to more than 10,000 Erlang/km^2) and flexibility in service creation must be added.

Intelligent network, IN, is a generic approach – independent of access – to open up the telephone networks to third-party service creation, to shorten introduction time for new services, and to support personalisation of services. A first global standard, ITU-TS Capability Set 1, CS1, is already agreed as a first step towards an architecture that meets the visionary definitions of personal telecommunication. The prime objective of the next step, CS2, is to support UPT.

The success of the PT endeavours hinges on the ability to define an architecture that supports the vision of personal telecommunication, and to strike a balance between a non-constraining reuse of investments already made and the introduction of new technologies in a way that minimises the impacts on existing networks. PT must integrate technologies used in PSTN, PLMN, ISDN and private networks to obtain low-cost transport, Fig. 1. IN technologies must be used to spread the control of service definitions to service providers and users. Technologies applied in the cellular networks – for managing the location of users – have to be used and further enhanced to support personalised services and multiple accesses as well. Microcell and picocell technologies, and hierarchial cell structures, are required to cater for high traffic density and to achieve low-cost terminals.

There is no single straightforward road to success. We will see an evolutionary process driven by different forces on an increasingly liberalised market, by the traditional PTTs, the mobile operators and the emerging "niche" operators competing with customised services. This process will not be controlled by long-term planning. The telecommunication industry may no longer afford to spend standardisation resources comparable to those spent on ISDN and GSM to reach the new objectives. The payback period might be too long. PT may be the first major enhancement of the global telephone network that is achieved by extensive reuse of existing investments and standards – driven forward by competition between the different players on the market, rewarding investments with a short payback period. It might also be the driving force towards a new access network infrastructure built on wireless standards.

In this process it is important to have a model that offers a way towards convergence in network development, whether the starting point is cellular or fixed networks, public or private.

Personal telecommunication – requirements

With some exceptions, current business can be characterised as the provisioning of standard services by *telecommunication organisations*, TO, to *subscribers*. The TO – Subscriber relation can be split up in a technical relation between the *network*

BOX A:

List of Acronyms

API	Application programming interface
AN	Access node
CCF	Call control function
CCS#7	Common channel signalling system No. 7
CT1	Cordless telephone standard No. 1
CT2	Cordless telephone standard No. 2
CT3	Cordless telephone standard No. 3
DECT	Digital European cordless telephone standard
DTMF	Dual tone multi frequency (touch tone) signalling
FSP	Flexible service profile
G-MSC	Gateway mobile switching centre
HLR	Home location register
IN	Intelligent network
INAP	CCS No. 7, intelligent network application part
ISDN	Integrated services digital network
ISUP	CCS No. 7, Integrated services user part
MSC	Mobile switching centre
NAP	Network access point
NP	Network provider
PLMN	Public land mobile network
PSTN	Public switched telephone network
PTT	Post telegraph and telephone
SCEF	Service creation environment function
SCEP	Service creation environment point
SCF	Service control function
SCP	Service control point
SDF	Service data function
SDP	Service data point
SIM	Subscriber identity module
SLEE	Service logic execution environment
SLP	Service logic program
SMF	Service management function
SMP	Service management point
SN	Serving node
SP	Service provider
SRF	Special resource function
IP	Intelligent peripheral
SSCP	Service switching and control point
SSF	Service switching function
SSP	Service switching point
TO	Telecommunication organisation
TUP	CCS No. 7, telephone user part
VLR	Visitor location register
V-MSC	Visitor mobile switching centre

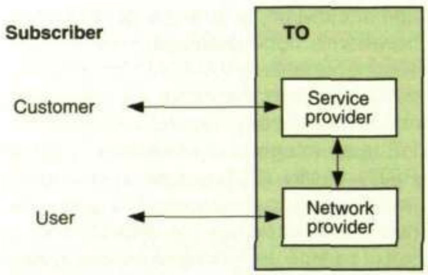


Fig. 2
A role model in service provisioning.
Each telecommunication organisation has at least two roles – that of service provider and that of network provider. Likewise, the subscriber has two roles: the customer procuring a service from an SP, and the user having terminal interfaces to an NP's network. The customer may sponsor a number of users, e.g. members of the family or the employees of the company

provider role and the *User* role, and in a commercial relation between the *service provider* role and the *customer* role, Fig. 2. The provider roles may develop into autonomous players, creating a scenario in which independent service providers develop service offerings to the customers and establish contracts with alternative network providers for the provisioning of these services.

Customer requirements

The subscriber – the customer – pays for the users he sponsors: the members of his family or the employees of his company. The customer may want to limit the authority of such users and possibly treat the traffic within the group of sponsored users differently from traffic to or from the group, irrespective of the users' location or whether they are stationary or on the move. The customer wants competitive prices with specified and consolidated bills and the possibility of choosing, at any time, the service provider that offers the best service.

User requirements

The user wants to be able to communicate:
- as a private or business person. This implies the requirement for numbering plans to ensure sufficient number space, since each role requires a number. It also requires a location management function in the network, capable of keeping track of each person's whereabouts. Multiple users may be registered on the same terminal, and use of differentiated alerting signals must be permitted
- with high probability of success. This implies the ability to determine actions to be taken in case of unavailability, such as redirection to other terminals, fax or voice mail
- with security and integrity. This implies requirements for registration and authentication whenever services are used
- at any convenient time. This implies the ability to define time windows for availability
- with selected persons only. This implies the ability to manage personal directories for both originating and terminating calls and to specify how each entry is to be treated
- in the way chosen. This implies the ability to choose the medium, such as voice, fax, data or mail, and to determine whether intermediate message storage is acceptable or whether a direct connection is required
- independent of terminal. This implies the possibility of using any available terminal (if permitted by the owner), and to be charged for the use of it. It also implies the ability to share a terminal with other persons
- when stationary or in motion. This implies the extension of the cellular network mobility to all access standards in all networks – without loss of feature transparency
- with comfort, limited only by the characteristics of the terminal and the access network. This implies user-friendliness and feature transparency.

Service provider requirements

The service provider wants maximal freedom in developing service packages that suit his target customer groups, meeting their requirements for customisation and

Box B

Standards for personal telecommunication

Intelligent network, IN
An ITU-TS/ETSI activity to define standards for IN. Capability Set 2, CS2, may be the first standard intentionally addressing enhanced versions of personal telecommunication, targeted for 1994 with implementations in the 1996 time frame.

Universal personal telecommunication, UPT
An ITU-TS/ETSI activity to define an architecture and service standards related to an access- and terminal-independent UPT number.

Universal mobile telecommunication system, UMTS
A long-term ETSI activity to define radio standards for future cellular services.

Future public land mobile telecommunication systems, FPLMTS
An ITU-TS activity to define standards for the third generation mobile systems for voice and non-voice services with universal roaming capabilities. The frequency range 1.7 - 2.69 GHz has been reserved for this purpose. Time target beyond 2000. UMTS and FPLMTS may well converge.

Advanced intelligent networks, AIN
BELLCORE is developing AIN standards for the Regional Bell Communication Companies in the US. The AIN Rel 0.2 version is under development in 1993 and 1994 as a candidate standard for PCS in the US, supporting both location management and remote call control as well as reuse of access port features in end offices.

Personal communication services, PCS
A broad effort in the US to determine standards and conditions for personal telecommunication with focus on low-cost, high-volume pocket telephones. High market pressure with field trials already in 1993.

Cordless telephone systems, CT1, CT2, CT3 and DECT
Radio standards for cordless telephones that can serve as normal telephones to the network.

American digital cellular, ADC
The designation used in this article for the US standard for the second generation digital cellular systems.

Groupe special mobile (global system for mobility), GSM
The ETSI standard for the second generation digital cellular systems.

Personal communication network, PCN
A popular commercial name for the use of GSM cellular technologies over and above the 1.5 GHz range.

Personal digital cellular, PDC
The Japanese standard for the second generation digital cellular systems.

personalisation. The service provider requests freedom to subcontract services from any network provider – services that are available anytime and anywhere, independent of access arrangement or terminals. The capability to quickly address new service markets and to create revenue by customising services in close cooperation with the customer is essential.

Network provider requirements
The network operator, or network provider, must be cost-competitive in every step of network development by expanding the network in a modular fashion, concurrently with the expanding market, introducing new technologies that offer lower costs from multiple sources, and allocating network service logic in the appropriate network nodes.

Automatic location management of users and their service profiles, which may be defined by multiple service providers, is a potential cost saver that replaces costly subscriber administration. The network provider needs to be in full control of the network to ensure integrity and quality. All efforts to open the network without sacrificing its integrity are merits.

Network architecture evolution

The communication model proposed is called the user agent model. This model – with its origin dating back to the early days of manual telephony – is used to characterise the different networks and technologies. In those days, operators provided both network and information services. The service level was only limited by the operators' ability to memorise their customers' whereabouts. Calling by name, recognition of voice, directory services and search for the wanted party when not at home, were embedded features of manual telephony. Personal telecommunication may be seen as a revival of this service but on a much larger scale, with higher availability, shorter delays and at affordable prices.

The user agent model
A user is an addressable entity in the network, identified by a directory number. Each user is represented in the network by

Fig. 3
An essential requirement for a personalised telephone service is to be able to screen incoming calls for certain periods. During these periods only calls from pre-selected users are put through – all other calls are diverted to an answering service or assistant

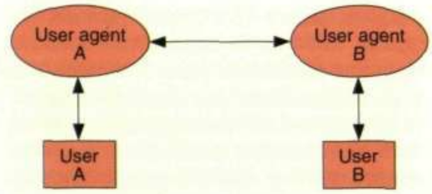


Fig. 4a
The user agent model.
Each user has an agent in the network. This agent communicates with its user for orders and with other user agents to effectuate the service

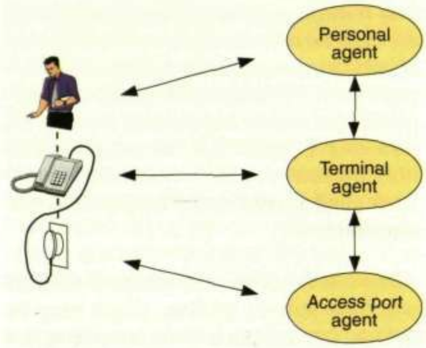


Fig. 4b
***Apart from the persons,* access ports and terminals also have their agents. This allows the user to access the network via different terminals and ports**

a user agent, Fig. 4a. This agent, which contains all service logic and service data related to the user, controls all communication sessions of the user, both with respect to location management and call management. The user may be an access port, a terminal or person – or rather a role of that person. The agents are access port agent, terminal agent and personal agent, respectively, Fig. 4b.

The users are connected to access nodes in the network, and the personal agents provide their services in serving nodes, Fig. 4c. Each user can call his own agent to change data in the service profile (service management), register the present location (location management) or request a communication session with another user (call management). Agents call each other using the home location number, which is an alias for the user number.

The different functions of the agent are executed or managed from the home location, the site in the network where the user is permanently registered. To meet cost and quality of service requirements, it must be possible to distribute the functions of the agent to a visited location at a serving node closer to users' present whereabouts, Fig. 5. *Location management* develops from user location management into location management of the user agent, its service management functions, its location management functions and its call management functions, Fig. 6. For cost reasons, *it is also necessary* that the access node and serving node can be co-located.

PSTN

The public telephone network, PSTN, is a heterogeneous network with different service definitions in different types of switch and different service offerings from different telecom operators. Normal access signalling is limited to hook signalling, dialling, ringing and acoustical signalling towards the user. Hook flashes are terminated in the access node. Only in the case of PBXs with direct inward dialling, the called party identity, too, is transferred to the terminating side.

Internally in the network, only the originating and terminating access port numbers can be transferred between originating and terminating access port agents. This

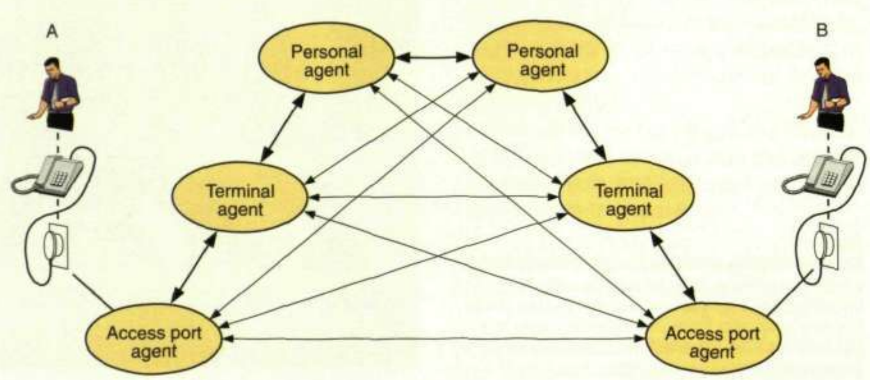


Fig. 4c
A description of user and agent relations for a call between two persons. Each person is using *a terminal that is semi-permanently or* dynamically connected to an access port

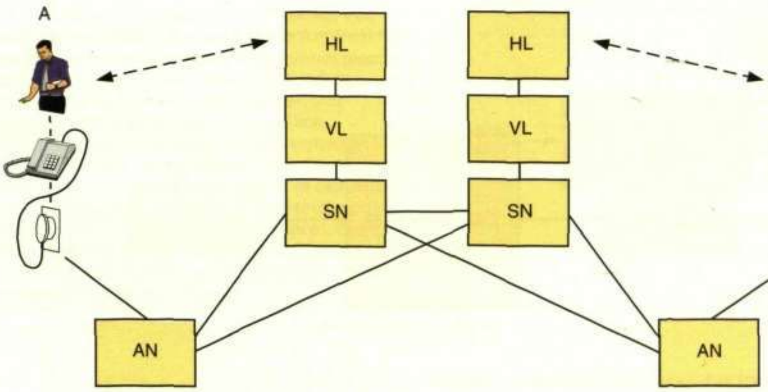


Fig. 5
A personal agent can control a call in a serving node, SN, by getting access to signalling information and control of switching functions. SN may be a separate node or integrated into the access node, AN. The user agent performs location management functions, service management functions, and call management functions. These functions are located in a home location, HL, and a visited location, VL

means that it is not possible to differentiate between terminals and users.

In PSTN, each access node is serving node, home location and, in the case of redirected calls, also the visited location for a partition of the directory number series, Fig. 7. Location management in modern SPC systems is performed by administrative command, semi-permanently associating the directory number, DN, with a physical port. The DN is normally referred to as the subscriber number, but obviously this is not always true. In the user agent model it is identical with the access port number.

Integrated services digital network, ISDN

ISDN is a monolithic network concept to integrate different teleservices and bearer services in one physical network. ISDN is characterised by its access having a signalling channel (D) common to two or 24/30 traffic-carrying channels (B) and by the use of a new common internodal CCS No. 7 application, ISUP. However, ISUP does not discriminate between terminal identity and user identity. A sub-addressing capability could possibly be used to address both a physical port and a certain terminal connected to this port.

Cellular networks for high speed mobility

In the second generation of standards for digital cellular systems, e. g. GSM, the base station controllers, BSC, act as access nodes. The mobile switching centre, MSC, with its integrated VLR (visitor location register) acts both as serving node and visited location with transparent signalling to the BSCs, Fig. 8. A separate location management mechanism has been developed to associate terminal identities with the geographical and physical addresses that may change when the terminals move. In GSM, the terminal receives its identity from the user's SIM card inserted in the terminal, without any association with its physical appearance in

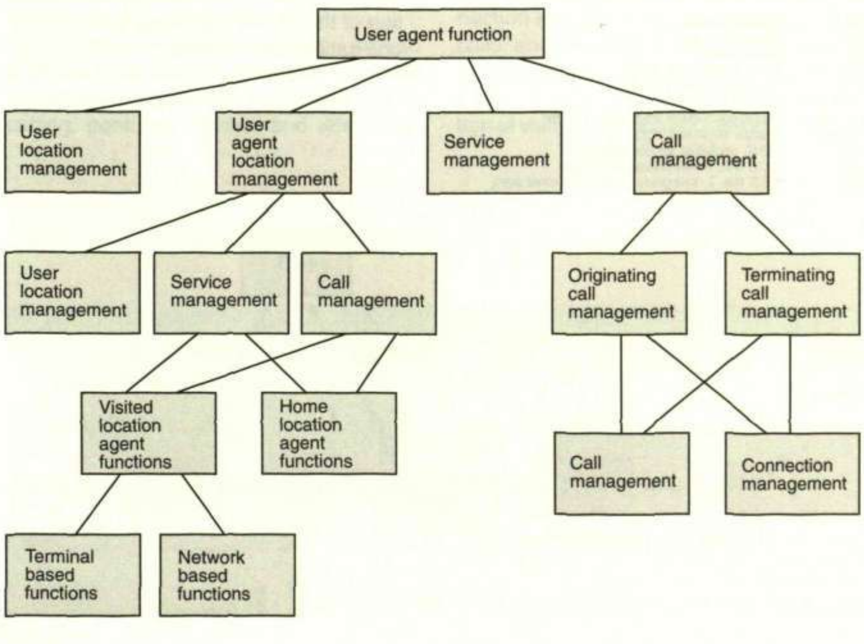


Fig. 6
The tree shows how the user agent functions might be conceptually differentiated into user location management, user agent location management, service management, and call management. The user location management functions keep track of the user, and the user agent location management functions keep track of the location of the different components of the user agent, whether at the home location, HL, at a temporary HL, or at a visited location, VL. The call management functions develop into originating and terminating call managers. Call management, in turn, may develop into separate call management and connection management

Fig. 7
PSTN and ISDN using the user agent model. The home location, HL, and visited location, VL, are determined by associating a directory number with an access port. It is not possible to differentiate between person, terminal and access ports. Routing in the network is performed by analysing segments of the directory number, which implicitly holds the identity of the serving node, SN, and the terminating access node, AN

CAS	Channel associated signalling
TUP	CSS No. 7, telephone user part
ISUP	CSS No. 7, integrated services user part

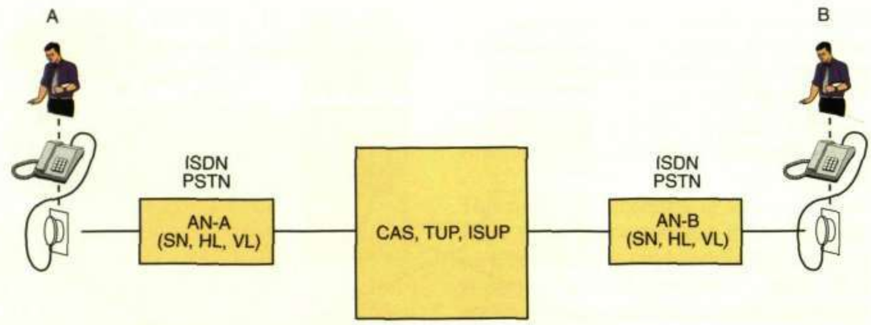
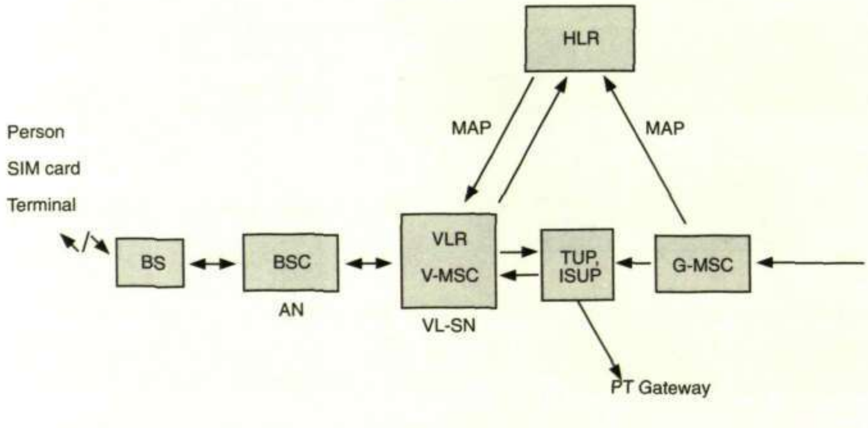


Fig. 8
The cellular network architecture. The terminal has an identity of its own or an identity given by an inserted personal SIM card. Registration and service management, as well as terminating call management based on user changeable data, are performed in the home location, HLR. Originating call management and terminating call management based on terminal status are handled by the visitor MSC, which is both visited location and serving node. Routing to a cellular terminal is made via a roaming number picked up via the G-MSC – HLR – V-MSC signalling

AN	Access node
SN	Serving node
BS	Base station
BSC	Base station controller
HLR	Home location register
VLR	Visitor location register
G-MSC	Gateway mobile switching centre
V-MSC	Visitor mobile switching centre
MAP	GSM, mobility application part
TUP	CCS No. 7, telephone user part
ISUP	CCS No. 7, integrated services user part



the network. An addressable entity, the home location Register, HLR, handles the terminal agent functions for a partition of the terminal number series. The HLR integrates a number of functions:

- location management of the call managers to secure that the flexible portion of the service profiles are currently updated in all visited locations (MSC + VLR) where the fixed portions of the profile are installed.

- assistance in call set-up to the terminal by forwarding call data to the VLR, and in return receiving the roaming number, RN, which is used to set up the connection for the call through PSTN. The RN is only used during call set-up, to associate the terminal number with the connection, thus circumventing the limitation of the PSTN signalling to carry only one called party number.

- direct communication with the terminals, using the MAP protocol to receive service management directives.

In a sense, the use of the personal SIM card unbundles the user from the terminal. However, current standards do not permit more than one user at a time to be registered at any one terminal.

The supplementary services are standardised. The majority of them, especially those using call state information, are implemented in the visited locations. Call forwarding services are performed by the HLR. The use of the same standard by a large number of operators gives feature transparency for the user in very large areas. GSM, for example, will cover the whole of Europe and a number of other countries. The large number of competing operators and vendors involved may make it difficult to agree on additions or amendments, which means that adaptations for personalisation may have to be made outside the standards.

Micro and pico cellular access networks for low speed mobility

To be able to support higher traffic densities, new practices for cell structuring and radio standards are required. Current mobile cellular practices support around 400 Erlang per km^2. The microcell and picocell practices adopted in the cordless standards CT2 and DECT can, for example, support up to 4,000 and 10,000 Erlang per km^2, respectively. However, with diminishing cell size, the capability to serve terminals moving at high speed is reduced. The solution is an hierarchical cell structure with mixed cell sizes, with an efficient location management ensuring that hand-

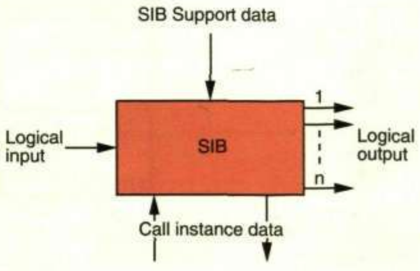


Fig. 9
The SIB is an elementary logical element in a service logic hiding the implementation from the programmer. When existing SIBs cannot meet a new requirement, new SIBs are defined. In Ericsson's IN products, the SIBs perform functions for e.g. analysis of signalling information, control of connection topology, interaction with the user, reading and writing of data, collection of call data, output of call data. Other SIBs are pure language elements, such as jump, go to subroutine, loop, handover. Each SIB is available in the service platform. Flexible Service Profiles, FSPs, are built by SIBs, referenced to by their names

overs are seamless and the frequency of handovers minimised.

Intelligent networks, IN

IN embodies the dream of the future unbundled network in which freedom is given to service providers and users to personalise the network services, independently of access, switch technology and network providers. An international consensus view on IN is described in the ITU-TS Recommendation Q.1200.

The unique characteristic of IN is that services are implemented on the IN service platform based on its service-independent building blocks, SIBs, and not directly in the network nodes, Fig. 9. Service logic can be designed using a service creation environment function, SCEF. The SIBs are made available to SCEF through a system-independent application programming interface, API.

So far, IN has been concentrated around a group of services – here called number services – for example, freephone, credit calling, personal number and televoting. Like PT, they provide service to numbers that are unbundled from the access ports in the access nodes. Any node in the network can be made a serving node by the addition of a service switching function, SSF, and a special resource function, SRF, both under control from a service control function, SCF, via a service-independent protocol interface. SCF is supported by a service data function, SDF, which may be physically unbundled from the node. Mapping of the functional entities into physical units is shown in Fig. 10.

The user agent is identified in SCF, by the calling or called party number, and called when an armed trigger point in the serving node is hit. Signalling data and call state data can be manipulated by the user agent. The SRFs are capable of in-band communication with the users, or with each other, to overcome limitations in the current signalling systems. Current IN standards assume that visited location and home location are collocated but possibly unbundled from the access node and serving node. Although the separation of access node and serving node functions reduce introduction costs, it results in potentially unwanted interactions between access port services and number-based services. An enhancement of the access node to a serving node is therefore required to ensure freedom in service design. An alternative is to add only two remotely changeable PT categories to the access nodes – one providing unconditional hot-line connection to the serving node for originating calls, and the other giving an unconditional call forward to the serving node for terminating calls. In the longer term, separation of visited and home location functions,

	SCEP	SMF	SDF	SCF	SSF	SRF	CCF
SCEP	X						
SMP		X					
SDP			X				
SCP			(X)	X			
IP						X	
SSP					X	(X)	X
SSCP			(X)	X	X	(X)	X
SN			(X)	X	X	(X)	X
NAP							X

Fig. 10
Mapping of IN functional entities into physical entities. The suffix F stands for function entity; P for physical entity

CC	Call control
SCE	Service creation environment
SM	Service management
SD	Service data
SC	Service control
SR	Special resource
SS	Service switching
SN	Service node
NAP	Network access point
IP	Intelligent peripheral

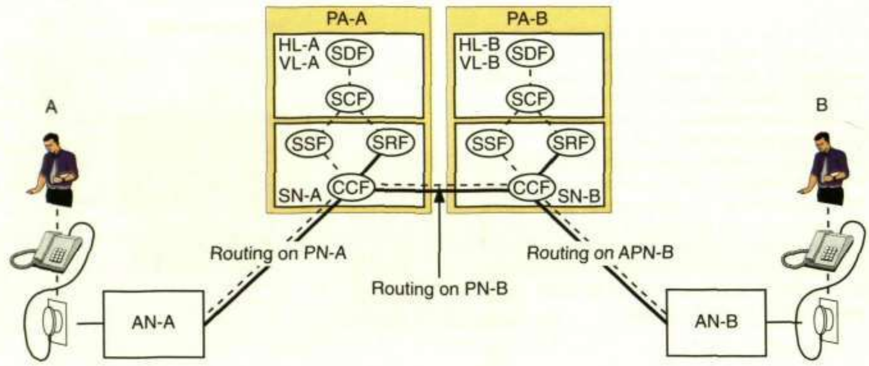


Fig. 11
An example of an IN implementation with serving nodes at the transit level. The serving nodes can be reached from any access node, such as a local switch in PSTN or ISDN or an MSC in PLMN. They can serve both PT and other number-based services. User identities and authentication information may be transferred in-band to the SRF or embedded in calling and called party number fields in the signalling systems. The personal agent has components in CCF (trigger point data), SCF (service logic) and SDF (service data). The IN platform components are either integrated into the access nodes or implemented in separate serving nodes

APN	Access port or terminal number
AN	Access node
CCF	Call control function
HL	Home location
PA	Personal agent
PN	Personal number
SN	Serving node
SCF	Service control function
SDF	Service data function
SSF	Service switching function
SRF	Special resource function
VL	Visited location

as in cellular networks, is required to reduce cost and improve capacity. Fig. 11 shows an example of an IN implementation.

Application programming interface, API.

The ITU-TS IN conceptual model, INCM, Fig. 12, has given rise to at least two approaches towards the application programming interface, API, Fig.14. One approach is to split the service logic into two parts: fixed logic and flexible logic. The SIBs are linked to form decision graphs that are called as subroutines by the fixed logic. The fixed logic is expressed in a standard programming language, such as C or C++, and requires compilation and loading into a standard execution environment. The flexible logic consists of exchangeable data only.

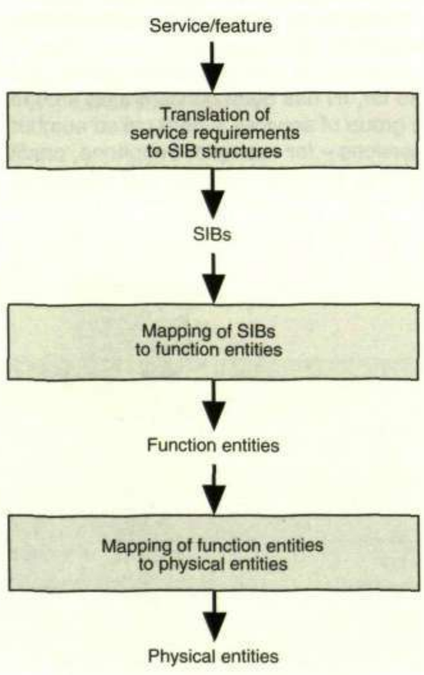


Fig. 12
The ITU-TS IN conceptual model, INCM. The service requirements are not directly translated into the network as has been the practice in all non-IN standards. Instead the requirements are translated into service platform elements called SIB (service independent building blocks). The SIBs are in turn implemented according to the ISDN three-stage model to become reusable capabilities and protocol elements

The second approach is to define a service API that gives full control over all aspects of the logic by combining SIBs with each other to achieve the desired function. Each SIB can be linked to any other SIB. Some SIBs perform a telecom function; others are only linking elements in the logic. All logic is expressed as data that describes which SIBs are to be used, how they are linked and what data each SIB is to use to perform its function. All implementation details are hidden to the service programmer. This is the approach taken in Ericsson's IN products. When the same data representation is used for all logic and data, the personal agents can be defined by means of flexible service profiles, FSP, Fig. 13. This gives a number of advantages:

- The different logics can be loaded and activated without service disruptions
- In case of a fault in the personal agent, only calls activating the faulty function will be affected.

Feature interaction

Feature interaction has been the major obstacle in the development of IN. The problem is that a feature is normally dependent on other features. There is a need to resolve such interactions, but no solution has yet been agreed on. As a rule of thumb, existing feature implementations are affected and have to be redesigned or completely blocked when new features are introduced. Two approaches to this prob-

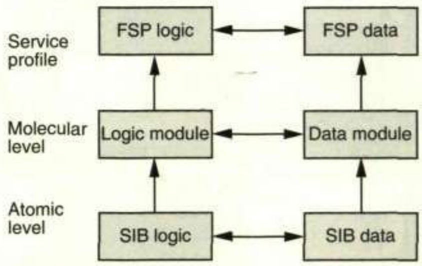


Fig. 13
FSP structure.
Each SIB has one elementary (atomic) logical component and one data component. The SIBs can be combined into molecular constructs with separate service logic and service data. Combinations of the molecular constructs give flexible service profiles, FSPs, for which logic and data may be managed separately

lem can be noted: the *network centric view* and the *user centric view.*

The traditional network centric view sees IN as a complement to other technologies in adding supplementary services to the existing repertoire. Interaction has been and continues to be the obstacle that prevents this view from being a realistic alternative. Each new supplementary service is composed of a fixed service logic part and, potentially, of a flexible logic part. Personalisation is limited to what can be achieved by combining a number of predefined supplementary services or features with each other. The adding of a new service may require long and costly development, not very different from the pre-IN experiences in PSTN, PLMN and ISDN. Designing of a new feature is not the issue, but rather how to integrate it with other features.

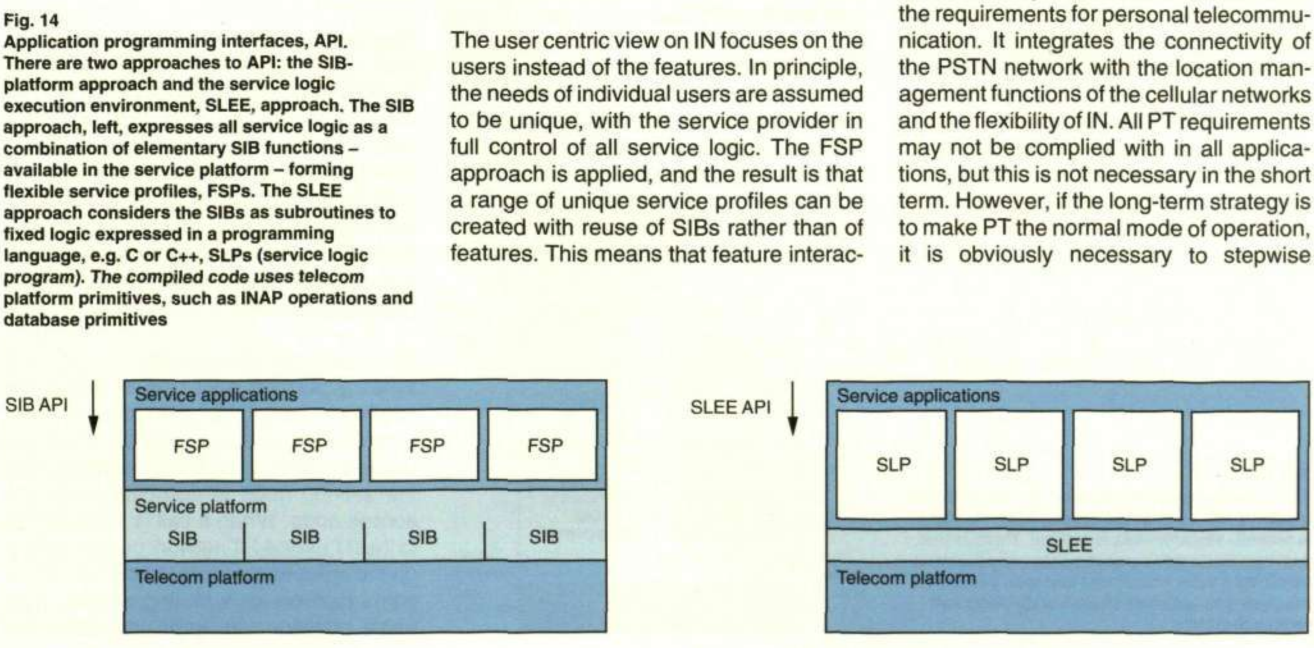
The user centric view on IN focuses on the users instead of the features. In principle, the needs of individual users are assumed to be unique, with the service provider in full control of all service logic. The FSP approach is applied, and the result is that a range of unique service profiles can be created with reuse of SIBs rather than of features. This means that feature interaction ceases to be a problem, since no individual features are implemented. The interaction between the SIBs constitutes the service logic. Interaction between service profiles is resolved through open signalling interfaces according to the half-call model. Before complete control can be provided from the stepwise developed IN platforms in an economically viable way, it might be necessary to use some of the existing supplementary services. It should then be borne in mind that this is a short cut that may result in interaction problems which sooner or later will have to be solved by an IN platform enhancement.

The target of the user centric view is to make the SIBs standardised to achieve both service independence and system and technology independence. When this is achieved, a SIB-based service profile can be executed on any compatible platform, whether it is a switch processor or a stand-alone personal computer or workstation. The old paradigm – to give the same features to all – is taken over by feature transparency for the individual, irrespective of access.

Personal telecommunication intelligent architecture

The architecture presented here should be seen as a way to reuse existing functions when new capabilities are added to meet the requirements for personal telecommunication. It integrates the connectivity of the PSTN network with the location management functions of the cellular networks and the flexibility of IN. All PT requirements may not be complied with in all applications, but this is not necessary in the short term. However, if the long-term strategy is to make PT the normal mode of operation, it is obviously necessary to stepwise

Fig. 14
Application programming interfaces, API.
There are two approaches to API: the SIB-platform approach and the service logic execution environment, SLEE, approach. The SIB approach, left, expresses all service logic as a combination of elementary SIB functions – available in the service platform – forming flexible service profiles, FSPs. The SLEE approach considers the SIBs as subroutines to fixed logic expressed in a programming language, e.g. C or C++, SLPs (service logic program). The compiled code uses telecom platform primitives, such as INAP operations and database primitives



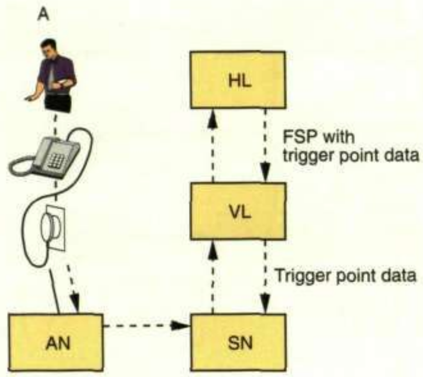


Fig. 15
The PT user has no designated terminal or access port. Instead the user identifies himself to the network from the terminal. The call is directed to a serving node, and a registration is made in the home location, HL. After authentication, the user's flexible service profile, FSP, is activated in the visited location, VL. The FSP also contains trigger point data which are fed to the serving node, SN. When a call hits an armed trigger, the corresponding FSP is activated

unbundle both service logic implementation from the design of the network node and the PT users from the terminals (or to unbundle the user with his terminal from the access). Clearly, such a strategy points out that PSTN and ISDN as well as PLMN and private networks will be affected by PT.

The personal agent
The personal agent is an application on the IN platform, with components in the SSF, SRF, SCF and SDF. It can be structured into a location manager, a call manager and a service manager – all of them allocated to the home location SCF/SDF of the user. The service manager communicates with its user for change of semipermanent data. The different components are expressed as FSPs.

Home locations
The location of a PT user can be determined by partitioning the PT number series into a number of home locations, in much the same manner as when allocating directory numbers to end offices in PSTN. This provides both for traditional routing algorithms, by which parts of the number are mapped to transit and local switch nodes, and for directory look-ups using the full number. Each home location must be sufficiently large to permit efficient management, and sufficiently small not to challenge technology and availability.

When a PT user registers, a connection is set up to the home location, authentication and service profile data is downloaded to the present visited location, and the corresponding data in a previous visited location is erased, Fig. 15.

Each call to or from a PT user means that the serving nodes involved have to exchange information with the home location. As the number of PT user increases and the PT network grows, the exchange of data will in itself constitute considerable data traffic. This traffic may be sufficiently heavy to motivate the introduction of a dedicated high-speed data network, possibly supplied by an independent network provider who also might provide the home location functions for a number of networks. Otherwise there is little room for using the IN protocols as interconnect standard between different providers.

Location management
When at PT user registers, the location manager allocates the call manager to the new visited location and erases the corresponding information in the previous visited location.

The call manager service profile
The call manager FSP has one part that controls originating calls and one part that controls terminating calls, Fig. 16. In the case of standardised service features, downloading is restricted to user-specific data. In the generic case, all service logic expressed as a function of SIBs is downloaded. International standards for the representation of such flexible service profiles have to be agreed.

Core network connectivity
When a call from a PT user is set up, the originating party's visited location is in full control of call management as in cellular networks. This is on the assumption that the serving node is collocated with the access node. When a call is to be set up to the PT user, a PT network gateway picks up the call and requests the called party's home location for a routing number. The home location, in turn, transmits the

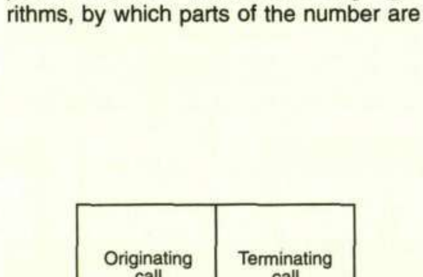


Fig. 16
The call manager FSP contains all logic and data unique to the user. In the special case the logic is shared, and only data is unique. There is one part that controls originating calls, one part that terminates calls, and a third common part that resolves interactions between originating and terminating calls

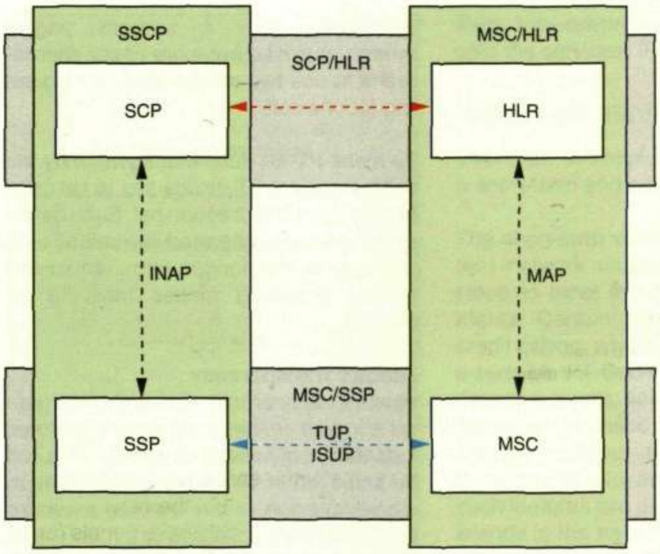


Fig. 17
Cellular network interworking options. When a cellular user is served by PT capabilities, interconnection can take place over TUP/ISUP interfaces between the mobile switching centre, MSC, and the serving node, SN (blue). Alternatively, SCF and HLR can interwork for exchange of user data (red). HLR can be integrated in MSC or be reached via the MAP protocol. The serving node can be an SSP interacting with an SCP over INAP or a combined SSP/SCP called SSCP. SCP and HLR may also be integrated in the same node, HLR/SCP. In the longer term, the MSC may become a PT serving node by integrating the SSF capability. Integration does not in itself change the service potential, which is determined by the TUP, ISUP, and INAP signalling capabilities

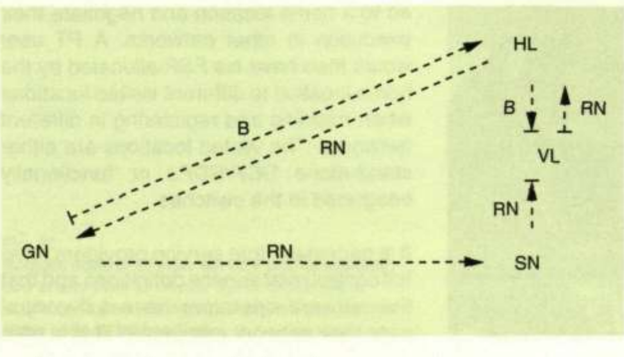
request together with call data to the terminating visited location to get a roaming number back. This number is returned to the gateway and used to set up the connection between the gateway and the terminating visited location, Fig. 18. The result is an association of call and connection. The purpose of this is to enable the use of a connection through PSTN despite its limited signalling capability. Only a small set of numbers is needed, since the setting-up time is very short. If the access node and serving node are separated, the same technique can be used to find the home location of the calling PT user.

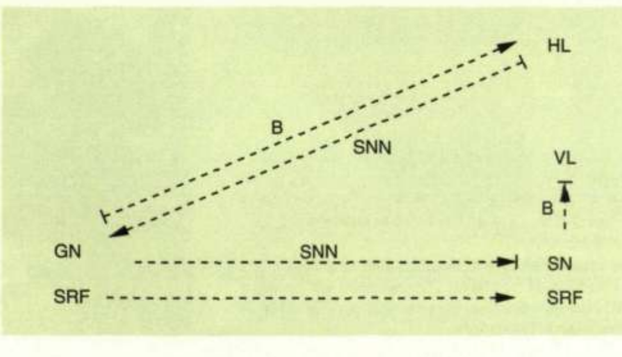
To enable origin-dependent analysis by the called PT user agent, it is necessary to transfer both personal and terminal identities as well as physical identities to the terminating visited location. To achieve this, current cellular practices have to be supplemented, Fig. 17. However, enhancement of TUP and ISUP with these capabilities is not to be expected in the near future.

PT terminals

To make PT possible over the PSTN, in-band signalling has to be used to convey all information required, including the PT identities. DTMF input is the natural choice with voice, tone or DTMF prompts. Voice-activated dialling, applying state-of-the-art voice recognition technology, is an emerging possibility. Full-duplex, in-band modem signalling between the terminal or PT device and the serving node may be the next step. An interesting alternative is to use a separate signalling channel,

Fig. 18
To enable the transfer of the user identity from the gateway node, GN, to the terminating serving node, SN, the cellular standard with roaming number, RN, can be used (left). Another alternative (right) is to set up a connection between GN and SN using the SN identity, SNN, and then transfer user identities in-band between special resource functions, SRFs, with high-speed modems





established over a wireless paging system. A similar but more costly alternative is to use two access ports for signalling for one call.

To make PT services manageable by the user, a special PT device has to be used for attachment to the terminal. Such devices can also be integrated in wireless telephones. In the longer term, enhanced access signalling modes have to be defined.

Feature transparency

Feature transparency means that the user will experience the same service features irrespective of network or access. This can be achieved in two ways: either through standardisation of the features available via the different accesses or through standardisation of service platforms common to all accesses. The first way is comparable to the GSM approach, which is one common standard to be used by a large number of operators. Applied to PT it would require one standard with uniform service features extended to all types of access – a proposal not very likely to be accepted, considering costs and lack of incentives to differentiation. The second way allows the user to experience personalised features as part of a personal service profile, limited only by the access and the capability of the network provider. IN technology can supply most features based on the service platforms, making reuse of access port features an exception. The strategy should rather be to block them when the access is used by a PT user.

Fig. 19
Personal telecommunication means a high probability of success. The network will locate and connect the person you wish to talk with anywhere on the globe

Network integrity for personal telecommunication in a multi-provider environment

IN is a new discipline which opens interfaces that have earlier been internal to the switching systems. These interfaces are complicated as compared with traditional open signalling interfaces, which loosely couple the network nodes with each other. Experience of multi-vendor interactions over the CCS#7 INAP (intelligent network application part) is gradually being gained. However, this interface has been developed for signalling within the networks and not for interconnection of different networks or operators. It follows that the operator of the switch also has to operate the service applications in the service control point, SCP, to be able to guarantee quality of service in terms of compliance with technical interface standards, availability of resources, and correctness of billing and other data.

INAP, in its present form, is not a suitable interface for interconnection of networks and service providers. A completely new type of interface has to be defined that ensures network integrity for both trigger point management in the switches and for service execution. However, a remaining obstacle with this approach is that the service provider cannot optimise his network by, for example, integrating execution of service logic in the switches.

A pragmatic view is to treat both SCF and SSF as parts of the network infrastructure under control of the network providers. Service creation and service management can be made available to service providers, who can have their FSPs downloaded to a home location and negotiate their execution in other networks. A PT user would then have his FSP allocated by the home location to different visited locations when roaming and registering in different networks. The visited locations are either stand-alone SCP/SDPs or functionally integrated in the switches.

It is necessary that service providers have full control over service definitions and that the network operators have full control over their network resources. This is best

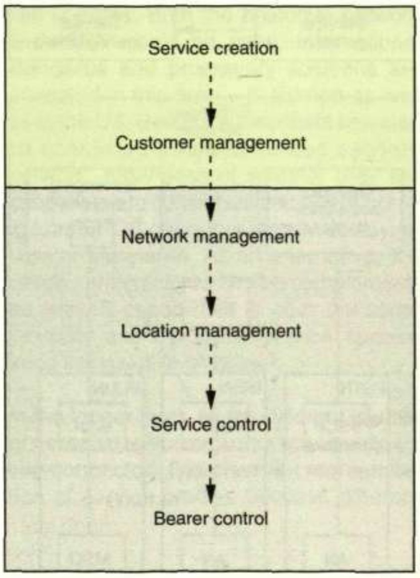


Fig. 20
Service provider scenario.
Service providers develop service profiles for their targeted customer segment. They negotiate contracts with different network providers for download of FSPs. The network providers verify the FSPs and execute the service

Handled by the service provider

Handled by the network provider(s)

assured by a split of their respective roles, such that the network providers receive and execute service logic and data from the service providers and respond with performance information. Service providers create and manage services for their customers and negotiate contracts with them. Network providers approve and execute the services, Fig. 20.

Implementation scenarios

There are two implementation scenarios: a short-term and a long-term scenario.

The short-term scenario extends the current network capabilities in incremental steps to meet the growing PT requirements. Certain new services, such as credit calling, which is a good example of a possible PT Service, are introduced on their own merits, service by service. Decisions can be based on short-term considerations, without taking any long-term investments into account. A number of such features are already implemented in islands in the networks – normally in end offices – or are under way. Examples include remote control of call forwarding, differentiated alerting, personal number, radio in the local loop, RLL. The cellular standards give wide-area mobility and are developed towards microcells and picocells. Terminal prices go down as a result of high volume and reduced cell sizes. Business cordless gives mobility in business communication networks. Cordless accesses may also be added in the cellular networks.

Fig. 21
Remote control of your service profile enables you at any time to inform the network about which calls you will accept and how you want other calls to be handled

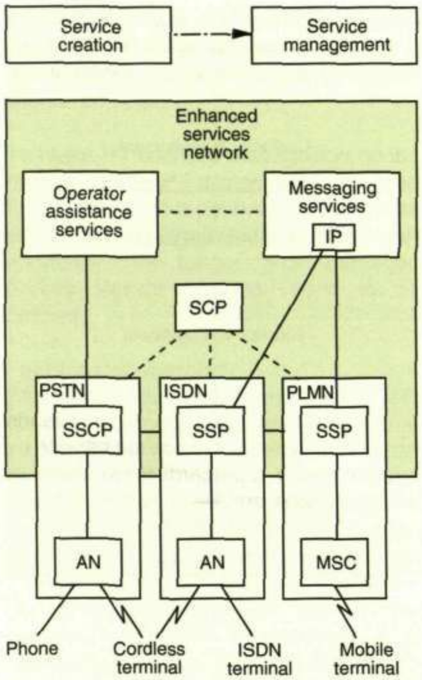


Fig. 22
Enhanced services network.
Personal telecommunication is one of many possible number-based services. Such capabilities are already offered by Ericsson's enhanced service network concept, providing IN solutions for number, messaging and operator services irrespective of access.

A single transit SSCP with combined serving node, visited location and home location functions can serve the whole network during an introductory phase. When more capacity is required, multiple SSCPs or SSPs with single or multiple SCPs can be introduced. The SCPs will then have both visited location and home location functions

—·— Control path
- - - - Speach/control path
——— Management data

In the long-term scenario, personal telecommunication is an architectural requirement imposed on the network in two dimensions. Firstly, to unbundle users from the accesses to achieve universal mobility. Secondly, to unbundle the service logic from the underlying network through the development and use of a uniform network service platform. Both requirements are necessary to generate new revenues, with the side effects of increased flexibility, technology independence and a concept for management of change with added incentives to competition in service provisioning.

The long-term scenario starts with the introduction of enhanced services network components at the transit level, Fig. 22. IN platforms in suitable configurations are installed for any number service and, specifically, PT services. Message store and forward equipment is connected to offer messaging services, such as fax and voice mail, and dedicated equipment is used to offer operator-assisted services. Any access – whether PSTN, ISDN or PLMN – can reach the PT user agents through normal routing functions. The users communicate with the personal agent, through voice prompts and DTMF input, using normal DTMF telephones or UPT devices with built-in intelligence and DTMF signalling capabilities. Home location and visited location functions are collocated. Home-based routing is applied, which means that calls to and from a PT user roaming in other networks are always routed via a serving node in the home network. Each user may have a personalised, call-controlling service profile, and it will also be possible to create user groups and virtual networks for scattered users. These features can be defined separately for each user and will always appear as being identical irrespective of access.

The use of cordless telephones that emulate the interface to the UPT serving node relieves the users of registrations and loca-

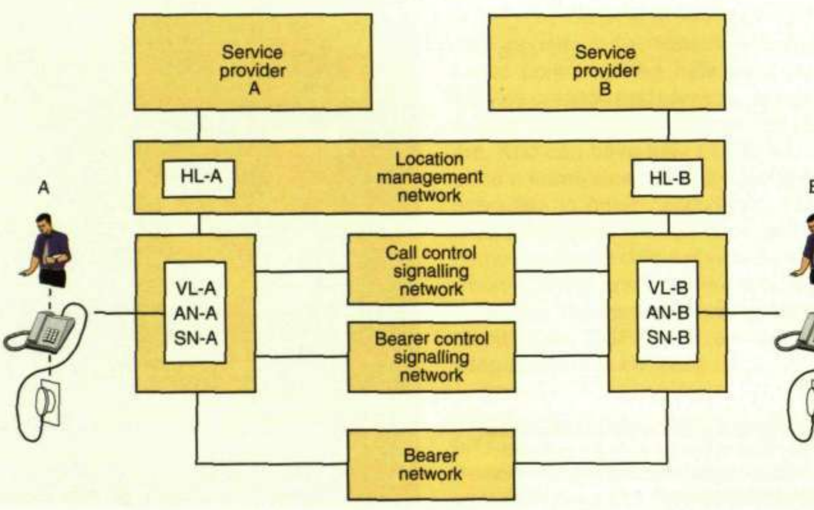


Fig. 23
In the longer term it may be necessary to move away from the home network based solutions by unbundling the visited location, VL, and home location, HL. SIBs and flexible service profiles, FSPs, need to be standardised. Download of FSPs from HL to VL in another network is possible. Separate high-speed data networks with directories for location management may develop. Access nodes and serving nodes are integrated. Access signalling will be the same independent of technology

tion updates. Both the analogue network and ISDN could be used. International standards and proprietary solutions are expected in this area – in Europe as well as in the US. Cellular subscribers can also be connected to the enhanced services network, regardless of whether they are subscribers to current cellular operators or to future PCS operators initially deploying cellular standards. As an alternative, the cellular architectures can be complemented with IN capabilities to offer the same flexibility and the same service appearance irrespective of access.

In the longer term, all the different islands of personal telecommunication have to be interconnected. Standards for representation of service profiles between different networks have to be developed, to make it possible to move away from the home-based solutions and thus reduce network costs, Fig. 23. Location management will become a major issue when every user is moving about and has a personal and flexible service profile assisting in managing the services, the terminal locations and the calls. Access and serving nodes have to be collocated for cost and performance reasons. Call control signalling and bearer control signalling may be separated when this is motivated.

All of this may be achieved the evolutionary way or through future UMTS and FPLMTS standards, which will also cover broadband, data, mail and electronic data interchange applications.

References

1. Söderberg, L.: *Architecture for Intelligent networks.* Ericsson Review *66* (1989):1 pp.13-22
2. Van Hal, P., van der Meer, J., Salah, N.: *The Service Script Interpreter, An Advanced Intelligent Network Platform.* Ericsson Review *67* (1990):1 pp. 2-12.
3. Ljungblom, F.: *A Service Management System for the Intelligent Network.* Ericsson Review *67* (1990):1 pp. 32-41.
4. ITU-TS Q.1200 series: *Recommendations on IN.*
5. Sundborg, J.: *Universal Personal Telecommunication (UPT) – Concept and Standardisation* Ericsson Review *70* (1993):4 pp. 140-155.

ERICSSON

ISSN 0014-0171

Telefonaktiebolaget LM Ericsson

Ljungföretagen, Örebro 1993