TELECOMMUNICATION SIGNALING

Telecommunication signaling is defined as a system that enables stored program control exchanges, network databases, and other nodes of the network to exchange messages related to call setup, call termination, supervision, information needed for distributed application processing, and network management information. Therefore signaling constitutes the command-control infrastructure of telecommunication networks.

Traditionally signaling is divided into two main types:

- *Subscriber loop signaling*—signaling between a subscriber terminal and the local exchange
- *Interexchange signaling*—signaling between exchanges which is also divided into two parts:
  - *Channel-associated signaling* (CAS)—signaling in the speech channel (in-band) or in a channel closely associated with the speech channel
  - *Common channel signaling* (CCS)—signaling in a channel totally separated from the speech channels and where this signaling channel is *common* for a large number of speech channels

In the context of telephony, signaling means passing information and instructions from one network node to another corresponding to the setting up, supervision, and termination of a telephone call. To initiate a call, a telephone subscriber (calling party) lifts the handset (hook) off. This action is a signal to the exchange that the subscriber wants to make a call. Then the exchange sends a dial tone back to the calling party, who then can start dialing a number. In due course the subscriber receives advice from the exchange about the status of the call, either a ringing tone, an engaged or busy tone signal, an equipment busy tone signal (congestion), or some other specialized tone.

Telephone signaling is also concerned with information transfer between exchanges (line and register signals). Register signals are used during the set-up phase of a call to transfer address and category information, and line signals are used during the entire time of a call to supervise the status of the line. The information contents in these signals are mainly the same as for the subscriber loop signals.

Until the mid-1960s, all such signaling was carried on or directly associated with the speech path. The following are examples of these CAS systems:

- One-voice frequency (decade pulsing)
- Two-voice frequencies (CCITT Signaling System No. 4)
- Multifrequency-pulsed [CCITT Signaling System R1 (No. 5)]
- Multifrequency-compelled (CCITT Signaling System R2)

The names of these signaling systems show that the most common way of transmitting the signals is in the form of pulses or tones (combination of tone frequencies). A common characteristic for this type of signaling is that there is one predefined signaling path for each speech channel.

The appearance of stored program control exchanges during the 1960s allowed introducing a common channel signaling system. In this new signaling concept, fast data links between the processors of the exchanges were used to carry all of the signaling, leaving the voice circuits to carry speech. In CCS a large number of circuits is handled by a few fast signaling data links. The signaling is performed in both directions, one signaling channel in each direction. The signaling information transferred is grouped into data packets called signal units. Besides the signaling information itself, there is also a need for speech circuit identification and address information and information for error control. Therefore stored program control exchanges together with the signaling links form a separate logical packet-switched signaling network.

There are two different standard systems for common channel signaling available. The first, CCITT Signaling System No. 6 (SS6), began developing at the end of the 1960s and is intended for use on analog lines, primarily for intercontinental traffic. The second, CCITT Signaling System No. 7 (SS7) was specified at the end of the 1970s.

SIGNALING SYSTEM NO. 7 (SS7)

SS7 is intended primarily for both national and international digital telephone networks, whose 64 kb/s transmission rate can be exploited. It may also be used on analog lines. During the 1980s, the demand for new types of services has increased. Therefore the SS7 has been developed to meet the signaling requirements from the following network types: the public switched telephone network; the integrated services digital network; the intelligent network; and the public land mobile network, especially digital mobile networks like GSM.

SS7 Network Components

A *signaling point* (SP) is a switching or processing node in a signaling network with the functions of SS7 implemented. A telephone exchange, functioning as a signaling point, must be of the SPC type. All SPs in SS7 network are identified by a unique code called a signaling point code (SPC).

A signaling link (SL) conveys the signaling messages between two SPs. Physically a SL consists of a signaling terminal at each end of the link and some kind of transmission media (normally a time slot in a PCM link) interconnecting these two terminals. A number of parallel signaling links that directly interconnect two SPs constitute a *signaling-link set.*

A signaling point, at which a message is received on one signaling link and then transferred to another link, without processing the contents of the message, is called a *signaling transfer point* (STP).

Signaling mode refers to the association between the path taken by a signaling message and the speech (or data) path to which the message refers. In the *associated mode* of signaling, the messages related to a call follow the same path as the speech between two adjacent SPs.

In the *quasi-associated mode* of signaling, the messages belonging to a call are conveyed over two or more link sets in tandem passing through one or more SPs other than those which are the origin and the destination of the messages. In this case the signaling messages follow a path other than the speech path.

A signaling point, at which a signaling message is generated, is called the *originating point.* A signaling point, to which a signaling message is destined, is called the *destination point.*

A *signaling route* is a predetermined path that a message takes through the signaling network between the origination point and the destination point. Therefore it consists of a succession of SP/STPs and interconnecting SLs.

All signaling routes that may be used between an origination point and a destination point by a message traversing the signaling network is the *signaling-route set* for that signaling relationship.

SS7 Functional Blocks

The SS7 includes a number of functional blocks. The *message transfer part* (MTP) functional block is a common transport system for reliable transfer of signaling messages over a signaling network. A number of different *user parts* (UP) are defined. Each contains the functions and procedures particular to a certain type of SS7 user. Examples of the UPs are signaling connection control part (SCCP), telephone user part (TUP) and ISDN user part (ISUP). The MTP transfers signaling messages between different UPs and is completely independent of the message content. The responsibility of the MTP is to convey signaling messages reliably from one UP to another. Figure 1 shows the MTP, SCCP, and other SS7 functional blocks.

The reliable transfer of messages between SS7 applications and user parts is provided by the *network services part* (NSP). It consists of the MTP and the SCCP. The MTP provides a connectionless message transfer function and the SCCP provides additional MTP functions for both connectionless and connection-oriented network services.

SS7 Signal Units

In the SS7 the signaling information is conveyed in *signal units* (SU), generally arranged in a form of packet-switched data communication. Figure 2 shows three main types of SU formats: *message signal unit* (MSU) contains the signaling information; *link status signal unit* (LSSU) is used for managing

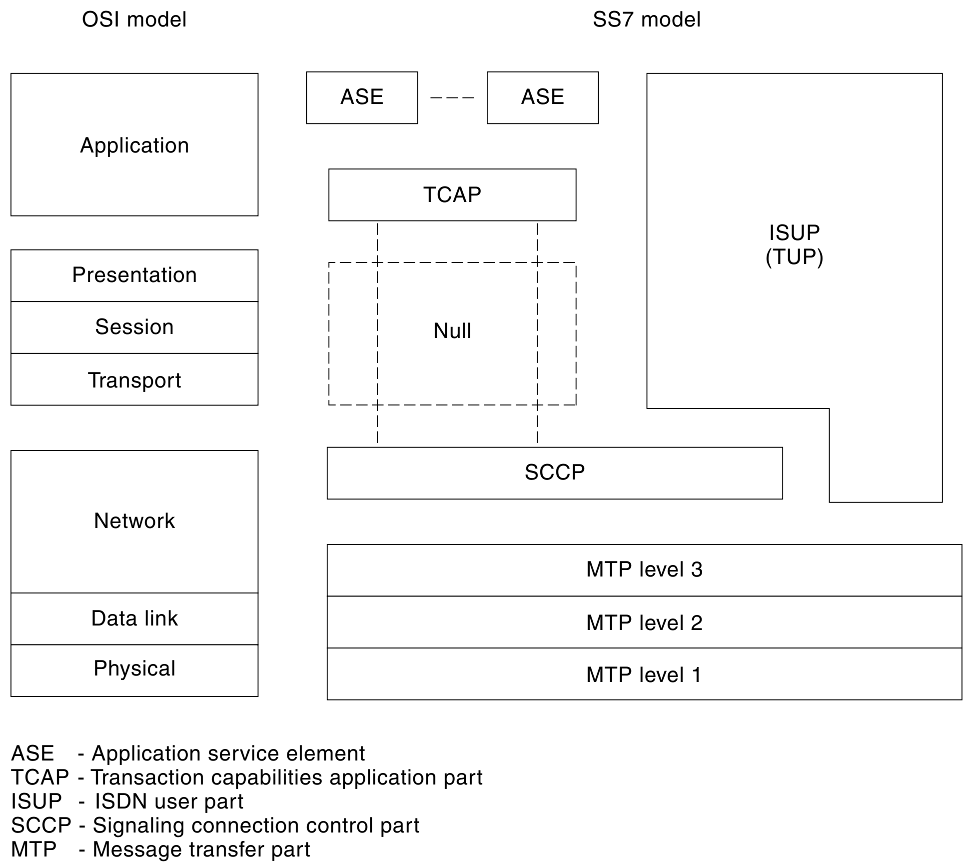


Figure 1. SS7 functional blocks.

signaling links; *fill-in signal unit* (FISU) is used as fill-in signals and for acknowledgement.

Each signal unit consists of a number of fields. The only field in MSU, *signaling information field* (SIF), has to do with the user parts. All of the other fields in all types of SUs contain information for the MTP. The SIF contains the signaling information from the user part and a label. Its length is less than or equal to 272 bytes.

The contents of the remaining fields in an MSU are as follows:

Service information octet (SIO) contains information about which user part the MSU belongs to.

Length indicator (LI) indicates the number of octets in the fields between the LI field and the CK field.

Check bits (CK) serves for detecting bit errors.

Forward sequence number (FSN), *backward sequence number* (BSN), *Forward Indicator Bit* (FIB), and *backward indicator bit* (BIB) fields are used in the error correction methods for sequence checking and for requesting retransmission.

Flag (F) indicates the beginning and the end of the signal unit.

Figure 3 shows the most common SS7 signals needed to set up a telephone call. They are as follows:

Initial address message (IAM) contains the main part of the called subscriber number (B) together with the calling subscriber category (À) and some other information

Subsequent address message (SAM) contains the remaining digits in the B-number

Address complete message (ACM) contains the status (free) of the B-subscriber

Answer, charge (ANC) usually starts the charging process

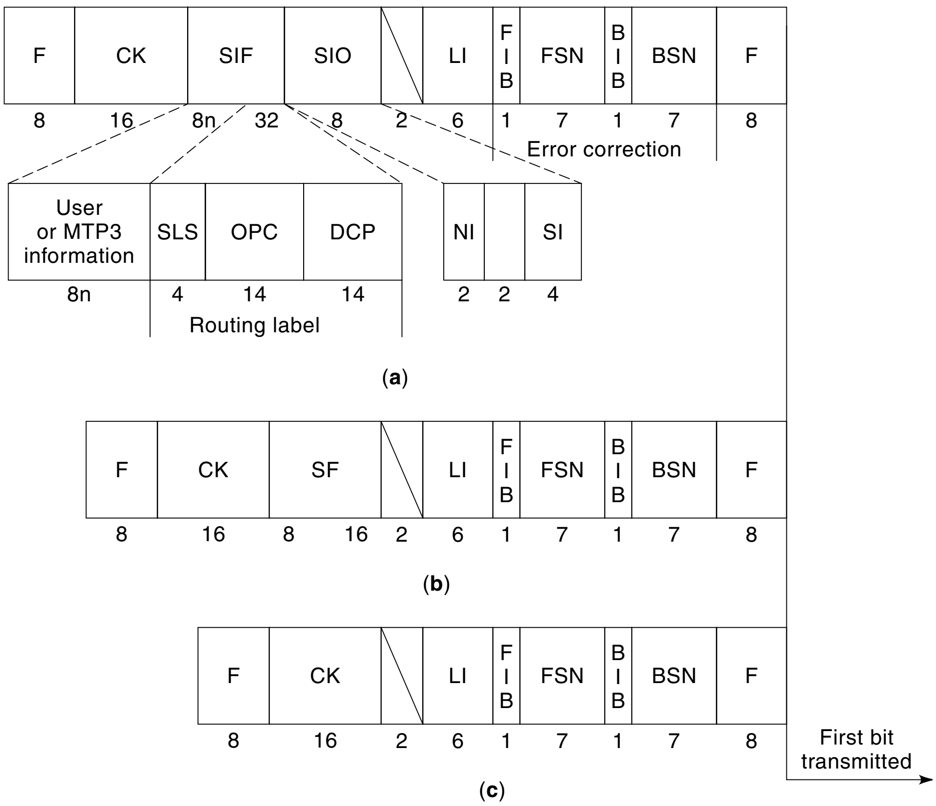
Clear forward (CLF) signals in the forward direction to terminate the call

Release guard (RLG) responds to the CLF signal and releases the circuits

Message Transfer Part

The overall purpose of MTP is to provide reliable transfer and delivery of signaling messages over the signaling network and to take necessary actions in response to system and network failures to ensure that reliable transfer is maintained. Figure 4 illustrates the functions of MTP and their relationship to the MTP users.

Signaling Data Link Functions (MTP1). A signaling data link is a bidirectional transmission path for signaling, consisting of two data channels operating together in opposite directions at the same data rate. It fully complies with the OSI's definition of physical level 1. A digital signaling data link is made up of digital transmission channels and their terminating equipment that has an interface to signaling terminals. The digital transmission channels may be derived from a digital multiplex stream having a frame structure as specified for PCM equipment or for data circuits. An analog signaling data



BIB: Backward indicator bit
BSN: Backward sequence number
CK: Check bits
F: Flag
FIB: Forward indicator bit
FSN: Forward sequence number
LI: Length indicator
SF: Status field
SIF: Signaling information field
SIO: Service information octet
SLS: Signaling link selection code
OPC: Originating point code
DPC: Destination point code
NI: Network indicator
SI: Service indicator

Figure 2. Signal unit formats.

link is made up of voice-frequency analog transmission channels and modems.

For digital signaling data links, the recommended bit rate for the ITU-T international standard is 64 kbit/s, and for the ANSI standard it is 56 kbit/s. Lower bit rates may be used, but the message delay requirements of the user parts must be taken into consideration. The minimum bit rate allowed for telephone call control applications is 4.8 kbit/s.

Signaling Link Functions (MTP2). The SS7 link functions show some similarity to typical data network bit-oriented link protocols, such as HDLC, but there are some important differences. They arise from the performance objectives of signaling that require the network to respond quickly to different kinds of failure events. The standard flag (01111110) is used to open and close signal units, and the standard ITU-T 16-bit cyclic redundancy check (CRC) check sum is used for error detec-

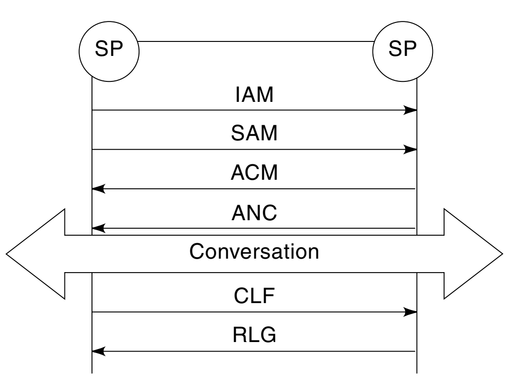


Figure 3. SS7 signals used to set up a telephone call.

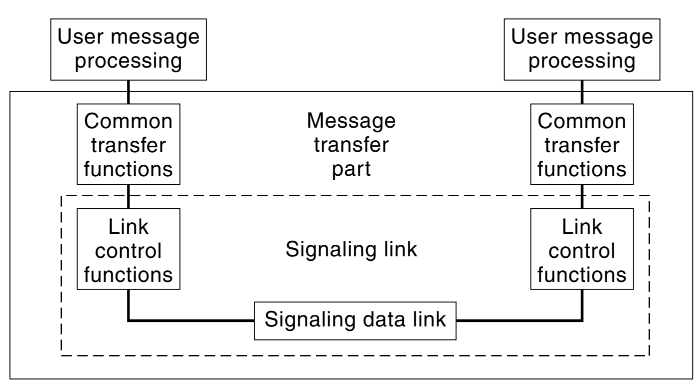


Figure 4. Message Transfer Part functions.

tion. When there is no message to be transmitted, FISUs are sent instead of sending flags as is done in other data link protocols. The reason for this is to allow for an error monitoring method so that faulty links are quickly detected and removed from service even when traffic is low.

Two forms of error correction are specified in the SL procedures. They are the basic error control method (BCM) and the preventive cyclic retransmission (PCR) method. In both methods only erroneous MSUs and LSSUs are corrected. Errors in FISUs are detected but not corrected. Both methods are also designed to avoid out of sequence and duplicated messages when error correction takes place. The PCR method is used when the propagative delay is large (e.g., with satellite transmission). With large propagative delays, the BCM method becomes inappropriate because the negative acknowledgment system causes message delays too long for erroneous MSUs. The drawback of PCR is that it allows for much less efficient bandwidth utilization than the BCM method. The maximum load level a link can be engineered for is much less with PCR.

The *BCM method* uses both positive and negative acknowledgments for error correction, as is done in many other protocols. If a negative acknowledgment is received, the transmitting terminal stops sending new MSUs, rolls back to the MSU received in error, and retransmits everything from that point before resuming transmission of new MSUs. Positive acknowledgments are used to indicate correct reception of MSUs and as an indication that the positively acknowledged buffered MSUs can be warded at the transmitting end. For sequence control, each signal unit is assigned forward and backward sequence numbers and forward and backward indicator bits (see Fig. 2). The sequence numbers are seven bits long. Therefore a maximum of 127 messages can be transmitted without receiving a positive acknowledgment.

The *PCR method* is a noncompelled, positive-acknowledgment, cyclic-retransmission, forward-error correction system. A copy of a transmitted MSU is contained in the buffer of transmitting terminal until a positive acknowledgment for that MSU is received. When there are no new MSUs to be sent, all MSUs not positively acknowledged are retransmitted. When the number of unacknowledged MSUs (either the number of messages or the number of octets) exceeds certain thresholds, it indicates that error correction is not getting done by cyclic retransmission. This would occur if the traffic level were high, thus causing a low retransmission rate. In this situation a forced retransmission procedure is invoked, new MSU transmission is stopped, and all unacknowledged MSUs are retransmitted. This forced retransmission continues until unacknowledged message and octet counts are below the lied threshold values.

Two types of *SL error rate monitoring* are provided. A signal unit error rate monitor (SUERM) is used while a SL is in service, and it provides the criteria for taking a SL out of service due to an excessive error rate. An alignment error rate monitor is used while a SL is in the proving state of the initial alignment procedure, and it provides the criteria for rejecting a SL for service during the initial alignment due to an excessive error rate.

The flow control procedure is initiated when congestion is detected at the receiving end of the SL. The congested receiving end notifies the transmitting end of its congestion with an LSSU indicating busy and withholds acknowledgments of all incoming signal units. This action stops the transmitting end from failing the link due to a time out on acknowledgments. However, if the congestion condition lasts too long (up to 6 s), the transmitting end will fail the link.

Signaling Network Functions (MTP3). The signaling network functions can be divided into two basic categories: signaling message handling and signaling network management. Figure 5 illustrates SS7 network functions.

Signaling message handling consists of message routing, discrimination, and distribution functions. These functions are performed at each SP in a signaling network and are based on the routing label and the SIO. The routing label consists of the destination point code (DPC), the origination point code (OPC), and the signaling link selection (SLS) field. The routing label is placed at the beginning of the SIF, and it is the common part of the label that is defined for each MTP user.

When a message comes from an MTP3 user, the choice of the particular SL on which it is to be sent is made by the *message routing function.* When a message is received from MTP2, the discrimination function is activated, and it determines if it is addressed to another SP or to itself based on the DPC in the message. If the received message is addressed to another SP and the receiving SP has STP function, the message is sent to the message routing function. If the received message is addressed to the receiving SP, the message distribution function is activated, and it delivers the message to the appropriate MTP user or MTP3 function based on the SI, a subfield of the SIO. Message routing is based on the DPC and the SLS in most cases. In some circumstances the SIO, or parts of it (the service indicator and network indicator), may need to be used.

Generally, more than one SL can be used to route a message to a particular DPC. The selection of the particular link to use is made using the SLS field. This procedure is called *load sharing.* Load sharing can be done over links in the same link set or over links not belonging to the same link set. The objective of load sharing is to keep the load as balanced as possible on the SLs within a link set. For messages that should be kept in sequence, the same SLS code is used so that the messages take the same path.

The purpose of the *signaling network management functions* is to provide reconfiguration of the signaling network in the case of SL or SP failures and to control traffic in the case of congestion or blockage. The objective is that, when a failure occurs, the reconfigurations are carried out so that messages are not lost, duplicated, or put out of sequence and that message delays do not become excessive. The signaling network management consists of three functions: signaling traffic management, signaling route management, and signaling link management. When a change in the status of a signaling link, signaling route, or signaling point occurs, then the following signaling network management functions are activated.

The *signaling traffic management function* is used to divert signaling traffic from unavailable SLs or routes to one or more alternative SLs or routes or to reduce traffic in the case of congestion without causing message loss or duplication. For SL unavailability events, a changeover procedure is used to divert signaling traffic to one or more alternative SLs. When an SL becomes available, a changeback procedure is used to

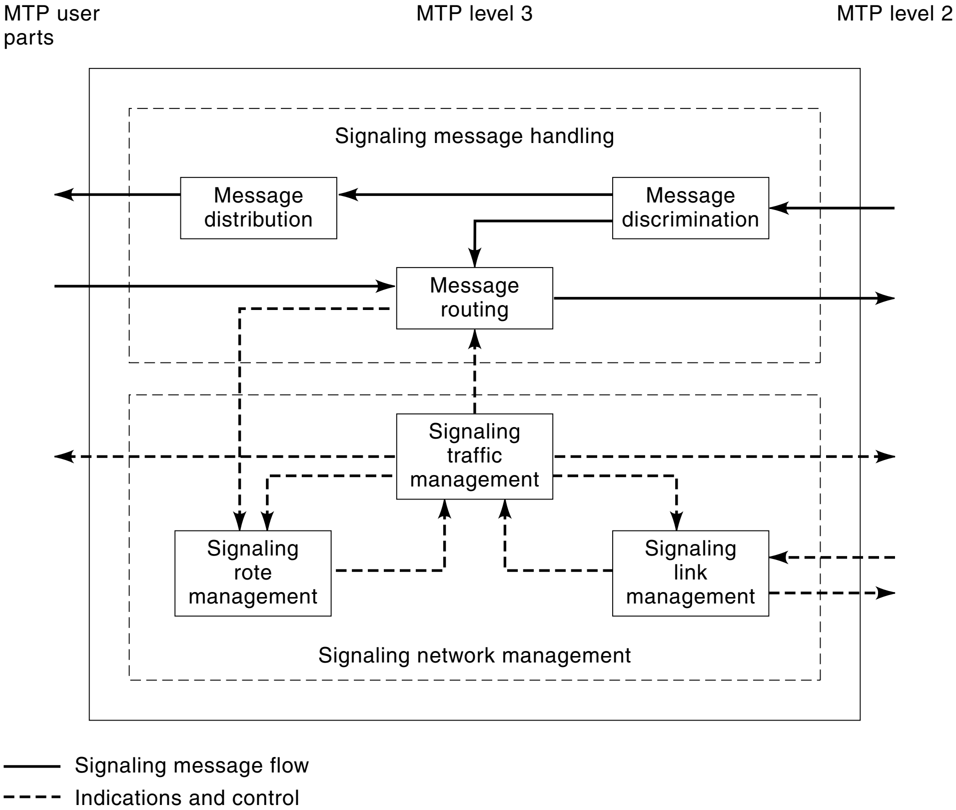


Figure 5. Signaling network functions.

reestablish signaling traffic on the initial SL made available. When signaling routes become unavailable, available forced rerouting and controlled rerouting procedures are used, respectively, to divert the traffic to alternative routes or to the route made available. Controlled rerouting is also used to divert traffic to an alternative route when a route becomes restricted. When a signaling point becomes available, the signaling point restart procedure is used to update the network routing status and control when signaling traffic is diverted to the point made available.

The *signaling route management function* is used to distribute information about the signaling network status to block or unblock signaling routes. The following procedures are defined to take care of different situations. The transfer-controlled procedure is performed at a signaling transfer point in the case of SL congestion. In this procedure, for every message received that a congestion priority less than the congestion level of the SL, a control message is sent to the OPC in the message. It tells the OPC to stop sending traffic that has a congestion priority less than the congestion level of the SL to the DPC in the message. In ANSI standards four congestion message priorities are used. In ITU-T standards only one is used. The transfer-prohibited procedure is performed at a STP to inform adjacent signaling points that they must no longer route to a DPC via that STP. This procedure would be invoked, for example, if the STP had no available routes to a particular destination. The transfer-restricted procedure is performed at a STP to inform adjacent SPs that if possible they should no longer route messages to a DPC via that STP. The transfer-allowed procedure is used to inform adjacent SPs that routing to a DPC through that STP is now normal. In ANSI standards, the previous procedures are also specified on a cluster basis, which significantly reduces the number of network management messages and related processing required when there is a cluster failure or recovery event. The signaling route-set procedure is used by the SPs receiving “transfer-prohibited” and “transfer-restricted” messages to recover the signaling route availability. Finally, in ANSI standards the signaling-route-set-congestion-test procedure is used to update the congestion status associated with a route toward a particular destination.

The *signaling link management function* is used to restore failed, activate new, and deactivate aligned SLs. There is a basic set of signaling link management procedures, and this set of procedures must be performed for any international or national signaling system. Two optional sets of signaling link management procedures are also provided, which allow for more efficient use of signaling equipment when signaling terminal devices have switched access to signaling data links. The basic set of procedures are signaling link activation, signaling link restoration, signaling link deactivation and signaling link set activation. The optional sets of procedures are procedures based on automatic allocation of signaling terminals and procedures based on automatic allocation of data links and signaling terminals.

The Signaling Connection Control Part

The addressing capability of MTP is limited to delivering a message to a node and using a four-bit service indicator (a subfield of the SIO) to distribute messages within the node. SCCP supplements this capability by providing an addressing capability that uses DPCs plus *subsystem numbers* (SSN). The SSN is local addressing information used by SCCP to identify each of the SCCP users at a node. Another addressing enhancement to MTP provided by SCCP is the ability to address messages with global titles, which are addresses, such as dialed digits, that do not explicitly contain information usable for routing by MTP. For global titles a translation capability is required in SCCP to translate the global title to a DPC plus SSN. This translation function can be performed at the originating point of the message or at another SP in the network.

In addition to enhanced addressing capability, SCCP provides the following four classes of service:

Class 0: Basic connectionless class

Class 1: Sequenced (MTP) connectionless class

Class 2: Basic connection-oriented class

Class 3: Flow control connection-oriented class

In class 0 service, higher layers pass a *network service data unit* (NSDU) to SCCP in the node of origin. It is transported to the SCCP function at the destination point in the user field of a unit data message. At the destination point it is delivered by SCCP to higher layers. The NSDUs are transported independently and may be delivered out of sequence. Thus, this class of service is connectionless.

In class 1, the features of class 0 are provided with an additional capability that allows a higher layer to indicate to SCCP that a particular stream of NSDUs should be delivered in sequence.

In class 2, setting up a temporary or permanent signaling connection (virtual circuits through the signaling network) performs a bidirectional transfer of NSDUs. Messages belonging to the same signaling connection are given the same SLS code to ensure sequencing. In addition, this class provides segmentation and reassembling capability. Therefore, if a NSDU is longer than 255 octets, it is split into multiple segments at the originating node, each NSDU segment is transported to the destination node in the data field of a data message, and at the destination node SCCP reassembles the original NSDU.

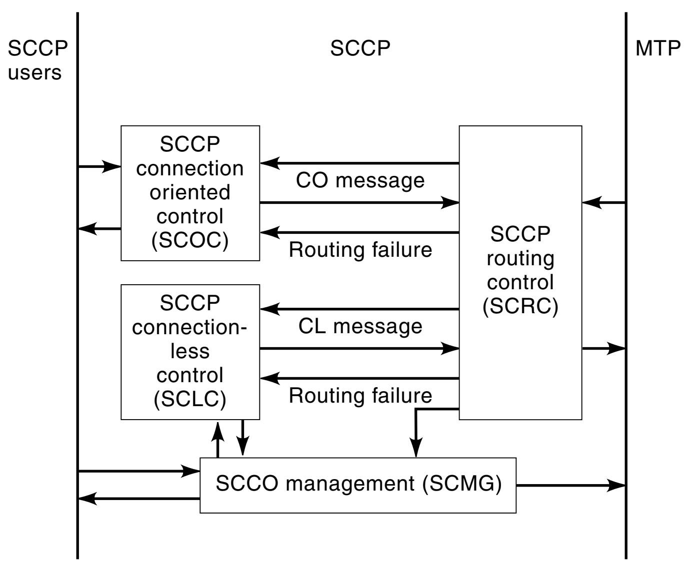


Figure 6. Signaling connection control part functional structure.

In class 3, the capabilities of class 2 are provided with the addition of flow control. The detection of message loss and out of sequence control are also provided. In such an event the signaling connection is reset and notification is given to the higher layers. The structure of SCCP, illustrated in Fig. 6, consists of four functional blocks. The SCCP connection-oriented control block controls the establishment and release of signaling connections and provides for data transfer on signaling connections. The SCCP connectionless control block provides for the connectionless transfer of data units. The SCCP management block provides capabilities beyond those of MTP to handle the congestion or failure of either the SCCP user or the signaling route to the SCCP user. The SCCP routing block takes messages received from MTP or other SCCP functional blocks and performs the necessary routing functions either to forward the message to MTP for transfer or to pass the message to other SCCP functional blocks.

BIBLIOGRAPHY

T. Russell, *Signaling System #7,* (2nd ed.), New York: McGraw-Hill, 1998.

Olli Martikainen
Helsinki University of Technology
Valeri Naoumov
Lappeenranta University of Technology
Konstantin Samouylov
People's Friendship University

TELECOMMUNICATIONS, SPACE. See Radiotelemetry.

TELECOMMUNICATION STANDARDS. See Telecommunication of Engineering Information.