S38.3115 Signaling Protocols – Lecture Notes

Lecture 4 – Integrated Services Digital Network (ISDN)

Contents[1]

Motivation for ISDN

The ISDN specification work started when it became obvious that future telephony systems would be digital back in the 1970's. CCITT released the so-called Red Book in 1984 and the Blue Book in 1988. The latter was technically sound enough for implementation although achieving interoperability remained a challenge. The result was that ISDN subscriber access started to be deployed in networks only in the 1990's. The specification work for message based trunk signaling had started earlier than the work on ISDN subscriber signaling. So, already the Red Book version of SS7 could be used for interoperable implementation.

[1] In 2015, this Chapter is recommended reading but is not included in the exam except for the section on Signaling Efficiency at the end.

The grand idea behind ISDN was that the network should be digital end to end and this would allow introducing new services to users and thus would help the operators to create new value and differentiate on the market.

Also, it was seen that digital switching is going to be superior to electromechanical switching because it will allow further automation of all operator business processes. To take full advantage of digital switching it was seen that native digital (= message based) signaling systems are needed. It was obvious that this will allow eliminating such special purpose hardware as DTMF receivers and senders, R2 multi-frequency receivers and senders[2] etc. The idea was also that message based signaling in combination with program controlled implementation of switching systems would allow implementing any number of services that would allow differentiation on the market.

First, digital signaling was introduced in trunk signaling (signaling systems 6 and later SS7). This is natural for two reasons: (1) digital switching first appeared in trunk switches and was introduced with a delay of several years in local exchanges. And (2) the changes in network nodes that are required to implement something as drastic as a completely new signaling system in a live network are limited to operator's own sites – no work gangs need to be sent to visit each and every subscriber. Note that a truck run to visit a single subscriber costs much more than an operator collects from that subscriber in a month.

On this course we will first talk about digital signaling for subscribers, i.e. the DSS1 signaling system, although it appeared later than SS7.

It is natural to split the operator's customers to two categories: (1) residential subscribers and (2) business customers. The first usually have one or two phone lines while the second often have a PBX or even a network of PBXs that need to be connected to a local exchange (or to several LEs). Small businesses (SOHO – small office home office) may also have a *key system* that in the analogue era was connected to the local exchange with a few analogue subscriber lines.

For local operators, the most long-term investment is the copper laid in ground and connecting users to its local exchanges. By introducing digital signaling on local loops, the operators thought that they will be able to add value to their investment into their copper plant and that they will leverage the value of their copper.

[2] (NB: senders are mentioned for completeness, they are easy to implement for a limited set of signals. One only needs to store the signals in a memory in a digital form such that the signals can be sent from the memory onto a PCM timeslot).

In practice it turned out that very few attractive subscriber services could be invented such that were impossible to implement with analogue subscriber lines and analogue subscriber signaling. Residential customers started to adopt ISDN in their homes more or less widely only, when Internet connections at home had become popular and modems for transferring data over analogue phone lines started to feel clumsy and slow.

On the other hand, DSS1 was widely adopted for the purpose of connecting PABXs to local exchanges and as a basis for specifying private network signaling systems. With DSS1, we had the first rich signaling system that was well specified and an international standard for corporate telephony services. For example, DDI stopped to be a problem. Reachability became easier. Clearly, easy reachability of company personnel is very valuable for businesses that actively communicate with their customers.

Why is ISDN important even today

Using DSS1 for PABX connectivity and for connecting special devices such as voice mail systems to exchanges is still common place. In addition, DSS1 has been used as the starting point for specifying GSM signaling. In terms of the circuit switched services 3G is quite similar to GSM.

So, even if DSS1 was not a huge success with residential customers, it became sort of a mother of many other signaling systems and many of its principles have been inherited to other more modern systems.

Method of specification used in ISDN

ISDN is specified in terms of *logical functions, reference points* and *interfaces*. Here, a logical function is a set of events an entity is able to understand, process and respond to. Reference points specify methods of communication of at least two logical functions. When a standard is implemented, one or several functions can be mapped to a single device (box). Interfaces are for the purpose of connecting devices that may be manufactured by different vendors. The idea is that each vendor implements its own box as it pleases for example using a combination of hardware and software. Because the implementation is left for the vendors, ISDN does not specify Application Programming Interfaces (APIs).

ISDN was specified in a series of CCITT books (red book 1984, blue book 1988, Euro ISDN 1992). The blue book was the first specification that was more or less accurate enough for implementation. However, only Euro ISDN gave enough detail for implementing supplementary services so that it started to make sense to produce user devices in reasonable volumes.

Digital transmission

Message based signaling can be understood by breaking the required functionality into layers and features on each layer. Messages are made of bits, these are organized into *frames* and finally inside the frames, we can find the messages that carry the telephony signaling information.

Line Coding

In order to transfer bits to a distance on a wire, an agreement what a logical one and a logical zero look like is needed. This agreement is called the *line code*. Physical limitations of the wired connection may dictate some properties of good line codes. Let us take an example.

The simplest code one can think of, for example on a physical copper connection like the local loop, would be to send a short pulse to indicate a logical one and zero voltage would indicate logical zero. In case of local loop this has the disadvantage that the code itself creates a constant current component (in a Fourier breakdown of the coded signal) that is dependent on what is being sent. This is not nice because the local loop is used for power supply and there may be things like transformers on the line card (something like this is needed for transforming a two wire connection to a four wire connection that is done on the line card).

To avoid creating a constant current component, a slightly more complex code can be used: the code is called Alternate Mark Inversion or AMI-code. The principle is that each logical one is sent with a different polarity than the previous one. It is also possible to use Reverse-AMI: logical zeros are mapped to pulses with alternate polarity and logical ones are sent with zero voltage.

The limitation with the AMI code is that a long string of zeros is difficult to count. The receiver needs to have a very accurate clock to be able to count the number of zeros that was sent. In reverse-AMI the difficulty comes with a long string of ones.

The difficulty of a possible long string of zeros is resolved in the HDB3 –code (more generally HDBn), see Figure 4.1. HDB comes from the words “high density bipolar 3”. This line code is used for example in the PCM system.

In HDB3, the forth consecutive zero is replaced by a pulse of the same polarity as the previous one violating the idea that there should be no constant current component and the basic encoding of logical zeros and ones. This violation needs to be balanced out. This is done by sending a balance pulse instead of the next zero.

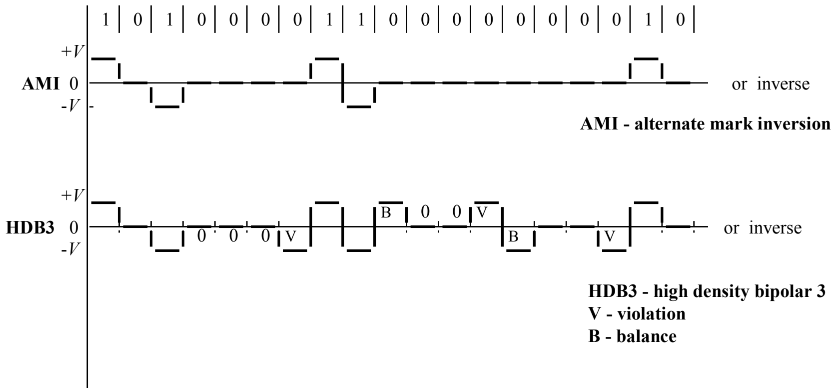

Figure 4.1 AMI and HDB3 line code.

Exercise: Make a coding table of 32 lines for HDB3 showing how all combinations of 5 bits are presented by positive and negative pulses and zeros.

Frame synchronization

In digital serial transmission, once we are able to detect logical zeros and ones, we need to spot bytes and strings of bytes. This requires that we detect the beginning and the end of a *string, we call it a frame*. Different transmission methods use different framing methods. The ISDN signaling channel uses a well-known method, inherited from HDLC (high level data link control).

HDLC framing is most suitable for transferring messages that are from tens to a few thousand bytes in length. The idea is to mark the beginning and the end of a frame with eight bits: “01111110”. This combination is called a *delimiter*. The immediate question is: what if the user wants to send 01111110? How can we make the distinction with the delimiter? The problem is solved by bit stuffing.

In HDLC, the sender, when it detects a series of 5 consecutive logical ones, always adds a zero onto the line. The receiver, always after 5 consecutive ones, removes the immediately following zero. If the receiver detects 6 consecutive ones, it determines that this is a delimiter.

Bit stuffing and HDLC framing are implemented on silicon and higher protocol layers need not worry about them. Processing HDLC framed messages is natural for computers: the contents are moved to memory and the message can be parsed using the CPU.

What if we need to send more than a few thousand bytes? We will discuss other frame structures later on this lecture.

ISDN Interface Types

ISDN Basic Rate Interface (BRI) is meant for residential customers and small offices. It uses a single copper pair of quality that was typical of countries where PSTN was widely spread. (Note that in offices today wiring is often CAT5 or better while the copper wiring in the ground is mostly of much poorer quality.)

Basic rate interface offers 2 data (or voice) channels of 64 kbit/s each, plus a packet transfer channel on 16 kbit/s. The last channel is mainly used for signaling but can also carry user data. So, a customer, using the existing copper wiring, could have 2 independent voice channels instead of one provided by PSTN. This turned out to be of limited attraction to the users. The added value came to fruition when faster connections to the Internet were needed. The data (or voice) channels are called B-channels (short for bearer) and the packet data and signaling channel is called the D-channel.

ISDN alternatively offers the *Primary Rate Interface* (PRI) that in practice makes use of PCM transmission. PRI is meant for connecting PABXs and servers like Voice Mail systems to exchanges. One PRI offers 30 data (or voice) channels and one signaling channel each on 64 kbit/s.

Both in BRI and PRI the D-channel is used for signaling related to calls that use any of the B-channels on that interface. The D-channel can also be used for transferring packet data. This possibility of transferring packet data has been made some use of in ISDN. The possibility later found a great success in GSM where signaling channels are used to carry short messages between users. One example of the use of packet data on D-channel is the always on-line connection to the Internet that was implemented using BRI. However, the use of this service in ISDN was short lived because ADSL came very quickly after its introduction.

From the ISDN exchanges point of view B-channels are continuous 64 kbit/s bit streams that are connected to its switching fabric, switched to PCM timeslots for wide area connectivity and allocated for calls based on signaling that takes place on the D-channel.

ISDN Reference Points

Reference points are standards connecting logical functions and potentially for connecting devices that come from different vendors or may be owned by different stakeholders (operators or customers).

BRI separates the home wiring (that is typically owned by the house owner) from the copper pair operated and owned by the operator. On this border, a Network Termination box (NT) is installed. Originally, the operators wanted to make this the demarcation point for responsibilities. Regulation forced the operators to offer BRI using different interfaces. Towards the user devices an

NT offers the S-interface (or reference point). On the copper wire towards the exchange we have the U-interface.

The relationship of a customer extension in the local exchange to user devices on a BRI is one to many (max 8). This means that the max of 8 user devices such as ISDN phones, fax-machines or PCs can be connected to a single BRI.

S-interface wiring uses 8 wires, two in each direction for data and four for power. Connectors in BRI are RJ45 (Note, the traditional phone jacks are RJ11). The S-interface was an international standard from the beginning. Its gross data speed is 192 kbit/s containing 2 × 64 kbit/s + 16 kbit/s = 144 kbit/s of user accessible data and 48 kbit/s of transmission overhead. Line coding on the S-interface is reverse AMI. The U-interface that became a standard was based on 2B1Q line coding on 160 kbit/s. Naturally, the same 2 data channels and the signaling channel with the total of 144 kbit/s need to be carried. On the U-interface, the reminder is the transmission overhead of 16 kbit/s.

On PRI, ISDN also provides the T-interface between the Network Termination 1 and Network Termination 2.

Between the Line Termination (LT) and the Exchange Terminal in the local exchange, ISDN specifies the optional V-interface. This was elaborated to V5.1 and V5.2 interfaces between subscriber multiplexers and LEs during the 1990's. These were however implemented by major wire-line switching system vendors in a slightly different manner and the goal that the operators had in mind was only partially achieved. The idea was that since ca. 70% of LE cost is in the access part, let's put a standard interface between the core switch and the access part and let the vendors compete in the two areas separately. Naturally, this was not particularly attractive to the vendors.

ISDN also talks about the R-interface. In fact this is a pseudonym for any legacy interface in the customer premises that can be mapped using a Terminal Adaptor (TA) box to the S-interface.

When ISDN was deployed, the boxes that the operator installed at homes were often actually combinations of the functions that are described in ISDN specifications. The home box could contain a TA for example for 2 analogue phone lines and an NT1 with 3 RJ45 jacks for S-interfaces.

Frame structure and synchronization on S-interface

Transmission on the S-interface is meant to be continuous. Line coding is AMI. To make sense of the bits, frames of 48 bits, 4000 times per second are sent in both directions (NT to user device and vice versa). The beginning of the frame is detected based on AMI code violation. Bit number 1 (F-bit) and number 14 (F_A) are zeros with a wrong polarity. There is a two bit time shift between the downstream frames (from NT to terminal) and upstream frames (TE to NT).

The code violation needed for frame synchronization is balanced using the L-bit.

Figure 4.2 shows the S-interface frame structure. The 48 bits in a frame have two samples for each B-channel (4000 times per second instead of 8000 times per second in PCM). Four of the bits in a frame are allocated to the D-channel. Next to D-channel bits, the NT echoes back downstream the D-channel bits it has received from a terminal. This uses another 4 bits (E-bits) on the downstream frame. The A-bit is used for Activation of the interface in the downstream direction. All this uses 1 + 1 + 4 × 8 + 2 × 4 + 1 = 43 bits. The rest is for achieving DC balance and future extensions.

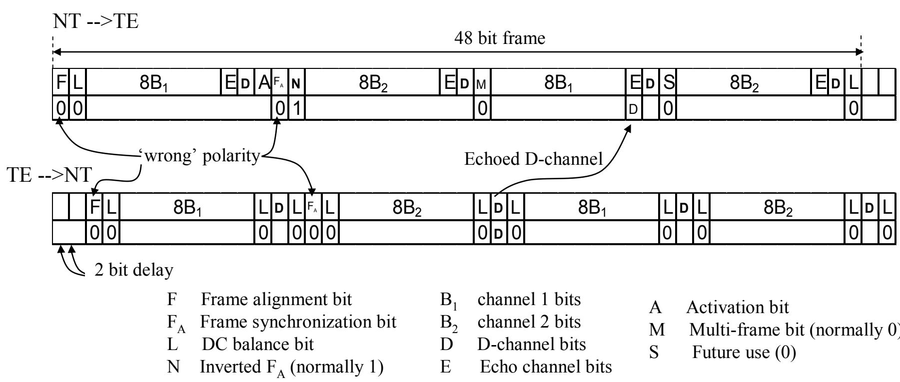

Figure 4.2: Frame structure on the S-interface.

Because up to 8 terminals can be connected to an S-interface, they need to compete for channel access. The echo channel shows to a sending TE that NT has received what it is sending. If the content does not match to what a TE is sending, signals from two TEs may have collided on the S-interface or if the problem persists, there may be a problem with the NT itself. For the case of collision a back-off algorithm is needed.

Let us summarize the discussion of framing in digital transmission so far. We are now familiar with HDLC framing for short sequences of bytes and two different framing systems for continuous digital signals: PCM framing and S-interface framing.

Power and power saving on BRI

For emergency calling it is required that besides data the LE sends power on the copper pair to the NT. It should be possible to make an emergency call without mains power supply (e.g. 240V AC) being present on the customer premises. Feeding subscriber lines that are not in use most of the time would be a burden and waste of energy. Therefore, the ISDN BRI interface usually goes to a power saving mode when it is not in use. To wake up from this mode, an activation procedure is required. Activation can be initiated either by a terminal or the LE.

A Terminal starts by sending INFO1 frames. When LE detects INFO1 or when it wishes to start the activation, it sends INFO2 frames. When Terminal detects INFO2 frames in the downstream, it starts sending active state INFO3 frames. LE acknowledges by starting to send INFO4 frames.

After this activation procedure, framing on the S-interface is in place and signaling on the D-channel can start based on the HDLC framing that we discussed earlier.

Layer 2 or the LAPD -protocol

The transfer of D-channel frames corresponds to layer 2 in Open Systems Interconnection (OSI). Layer 2 frames can carry either signaling messages of layer 3 or packet data. Layer 2 protocol in ISDN is called LAPD for "Link Access Protocol on D-channel" and it is specified in Q.920 –Q.921. Layer 2 is concerned with a reliable transfer of data between a set of terminals and the LE. Layer 2 does not understand about calls. Calls are the concern for layer 3, the DSS1 signaling that is specified in Q.930-Q.931.

To be more exact, LAPD serves for the purpose of reliable transfer of chunks of data between many terminals and many layer 3 entities in the LE.

The first thing one almost always needs to understand about a protocol is how the communicating parties are identified and thus who can communicate with whom.

In case of LAPD, the parties are identified with a Data Link Connection ID (DLCI) field that is carried in each frame. DLCI normally takes 16 bits after the delimiter marking the beginning of the frame. The DLCI can be extended to larger values if needed. DLCI has two subfields: SAPI or Service Access Point Id and TEI for Terminal Endpoint Id. SAPI identifies the layer 3 or management entity to which the frame is addressed. TEI values can be dynamically allocated to terminals in a BRI. LAPD has procedures for the dynamic allocation of TEI values.

Before any higher level (Q.931) functions can be performed, each ISDN device must be assigned at least one unique TEI value. These numbers can be pre-assigned (TEIs 0-63), or dynamically assigned (TEIs 64-126). Most TEI assignment is done dynamically, using the TEI management protocol that is part of LAPD. Using dynamic allocation eases configuration.

Next after the address field, LAPD has a control field of two bytes. This is used for identifying the types of frames (I, S and U –frames) and carrying frame numbers. LAPD supports both an acknowledged mode of data transfer with I-frames and unacknowledged mode with IU-frames. In the acknowledged mode LAPD supports *windowing*: a certain number of frames can be sent before an acknowledgement MUST be received prior to sending more. By frame numbering reliable transfer of data can be achieved: the receiver can see if a packet is missing and respond to the sender accordingly.

Each layer in a protocol stack usually needs some management functions. This is called layer management. An example is the TEI management procedures we already mentioned.

After the address and control bytes, a LAPD frame has room for the payload. The frame ends with a two-byte checksum and the ending delimiter.

Telecommunications and Bearer Services

In ISDN the term "*telecommunication service*" refers to functionality that covers all OSI layers. The term *bearer service* refers to data transfer on layers 1 up to 3. In such bearer services as audio, speech or transparent 64 kbit/s data only layer 1 is seen by a local exchange.

- ✓ Circuit switched bearer services supported by ISDN are:
  - › Speech
  - › 3,1 kHz audio
  - › 7 kHz audio
  - › transparent 64 kbit/s.
- ✓ Packet switched bearer services (on D-channel) in ISDN are:
  - › virtual call and permanent virtual connection,
  - › connectionless packet switched service on the D-channel,
  - › user-to-user signaling information.

Telecommunication services make use of some bearer service. Telecommunication services are further broken down into *basic services* and *supplementary services*. An example of a basic service is a telephone call. Supplementary services cannot exist without a basic service.

Q.931 ISDN signaling

Q.931 is the layer 3 protocol used on Digital Signaling System Number 1 (DSS1). It is carried in the payload of LAPD.

Because D-channel is separate from the voice channels, DSS1 is *out-of-band* and can continue independent of the use of voice channels. Encoding of Q.931 is *binary* (not for example text) and the style of encoding is that information in the signaling message is grouped into *information elements*. Each information element has a *type field* that identifies the kind of information that may be contained in the element. If the element is of variable length, a length field gives the number of bytes in the element. It follows from this coding style that the required message parser is Q.931 specific.

Each message has a message type field. This gives an idea about the main function or state transition that the receiver should execute on reception.

Identification in Q.931

Q.931 uses telephone numbers for *identifying subscribers*. The layer 3 entities related to a call are identified by *Call References* that are arbitrarily chosen by the communicating parties (Terminal or LE). A signaling message may *identify the B-channel* that is allocated for a call in an information element that is carried inside the signaling message.

All this means that Q.931 signaling is also logically separated from the voice channels. Signaling entities can be created and they can live independent of the allocation of voice channels.

ISDN BRI supports the max of 8 terminals in a single S-interface. They can all use the same telephone number or each one can have its own telephone number. The first option leads to a need to identify compatibility of the terminal with the service that is being requested. For this purpose Q.931 supports *Lower layer compatibility* information and *High layer compatibility* information elements during the call setup procedure. Compatibility info can for example define that a G4 fax machine is requested and that other terminals in the S-interface should not react to an incoming call. Actually, this violates the layering principles that are assumed in OSI. Compatibility relates to higher layers than layer 3. Nevertheless, the information about compatibility is carried in Q.931 that is supposed to be a layer 3 protocol. The advantage is that if there is no compatible device connected to the concerned S-interface, no charges will be incurred to the calling subscriber. Instead the call will fail with a minimum cost to the operator. The alternative would be to let the user devices exchange whatever information they wish prior to setting up a call (for example of less than N bytes). This, however, did not meet the operator's charging requirements. Letting users communicate in any way prior to setting up a call and starting charging just does not follow the telephone operator principles.

Signaling modes

Q.931 supports two modes of signaling: the *functional mode* and the *stimulus mode*. In functional mode information is encoded in service specific information elements. This sounds very logical but unfortunately, it means that software in both ends of communications needs to be updated if new functional information elements are needed for new services. There is no such thing in ISDN as software download to Terminals from the network. This leads to cumbersome updating of Terminals and for example JAVA was invented much later than ISDN.

Stimulus mode means that a signaling message carries the keying sequence on the phone to the LE. It is possible to program new keying sequences on programmable phones (soft keys) and implement new services just by updating the LE.

Q.931 supports *overlap sending*. This means that the dialing sequence can be as it was for PSTN phones: pick up the receiver, hear dial tone, push buttons for dialing and wait for the ringing tone. The opposite of overlap sending is *en-block sending*. In this mode all dialed digits are sent in the first signaling message. A special key is needed on the phone to complete the dialing sequence. The same mode is always used in mobile dialing: one first selects all digits and then pushes the call button. If the numbering plan allows only fixed length subscriber numbers, it is feasible to collect the dialed digits in the phone even in fixed telephony and send them all at once in the first signaling message. On the other hand, in countries with variable length telephone numbering (like Finland), overlap sending speeds up connection setup and avoids storing information about the dialing plan onto phones.

Call setup and release

DSS 1/Q.931 signaling is symmetric. This means that *originating call signaling* and *terminating call signaling* use the same messages, just the direction is (LE –terminal) is reversed. Symmetry applies to both call setup and release. We show the signaling flow in Figure 4.3.

The first message for call setup is the SETUP message. SETUP message contains at least the bearer service (= is this a data call or audio call). In practice it is a good idea to include as many digits in the SETUP message as there are in the shortest directory number in the local dialing plan (4 or 5 in the Finnish ISDN network e.g. 47001 for Aalto University), so the originating exchange can attempt routing immediately having received the SETUP message.

SETUP_ACK acknowledges the reception of SETUP. It is more useful on the terminating side – tells that the kind of device that can support the bearer and telecommunication service requested has been found (e.g. a G4 fax machine). INFO messages support overlap sending – the result is that the routeing can be done through the network with the minimal number of digits (in Finland usually 3). In practice INFO messages are useful only in the TE to LE direction for the TE to send each new dialed digit in a separate message. This will lead to the fastest possible call setup: often the ringing tone is heard immediately after pushing the last button on the phone.

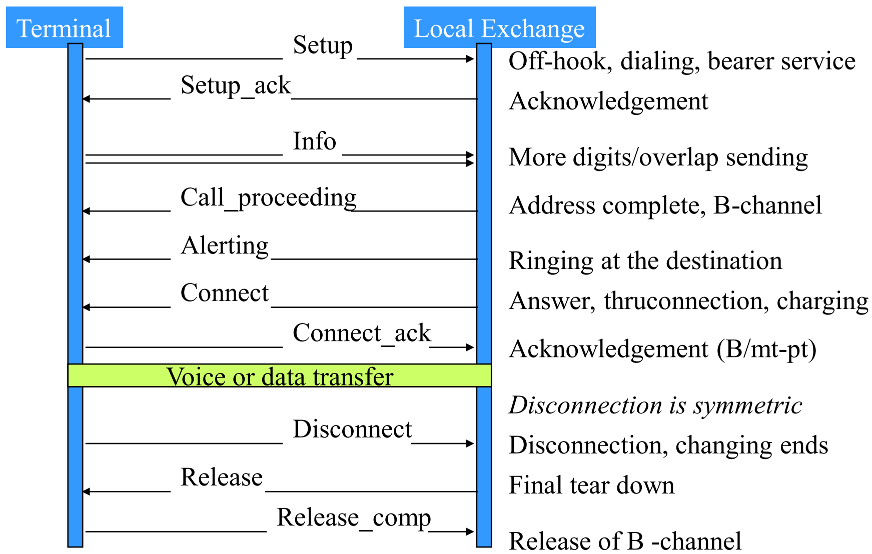

Figure 4.3: DSS1 Signaling Flow for Call Set-up and Release

NB1: In case there is no number portability (NP), the last digits (in Finland digits 4…8) are needed only at the terminating exchange, the rest of the exchanges do not even look at the last digits. If NP is allowed in the number block, all digits need to be received prior to routeing decision at the originating exchange.

NB2: There is no such thing as number complete indication from the terminal, instead the network deduces when all digits have been received by executing number analysis after each digit or at times instructed by former routing information retrieved based on earlier digits received for the call.

CALL_PROCEEDING in the backward direction – tells that no more digits will be needed even at the terminating exchange, it also tells that at least the network is not busy (the B-subs can still be busy).

ALERTING in the backward direction tells that the phone at B-subscriber (called party) is ringing. The D-channel message is accompanied by a *ringing tone* from the terminating exchange to the caller on the audio channel. This tells that the audio channel is clear at least in the backward direction from the terminating exchange to the caller and that the B-subscriber is not busy.

CONNECT message starts charging and all exchanges through-connect the B-channel in both directions. Charging data collection usually takes place at the originating LE. B-subscriber has answered the call.

CONNECT ACK is really necessary at the terminating side: tells which of max of 8 devices that can be connected to the S-interface the call has been awarded (imagine two or more phones being picked up almost at the same time each sending a CONNECT).

DISCONNECT – charging stops and the B-ch connection is torn down. One of the subscribers has gone on-hook.

RELEASE tells that DISCONNECT was received. RELEASE COMPLETE confirms the reception of RELEASE and tells that after this the B-channel can be used for another call on the subscriber interface. For call tear down in exceptional cases even RELEASE and RELEASE COMPLETE are enough.

One should also note that the call can be released any time during call setup if the caller goes on-hook. Instead of CONNECT, there may be a DISCONNECT in the backward direction etc.

If one takes a look at the call setup procedure, it becomes clear that Q.931 does not follow the client-server model. Instead the model is that of communicating finite state machines.

Connecting PABXs to ISDN networks

We already mentioned that DSS1 has been widely used for connecting PABXs to ISDN networks and as a basis for private network signaling systems. The move to digital signaling for PABX access started already prior to CCITT (predecessor of ITU-T) specifications of DSS1 being ready. In the UK pre-ISDN standards for DASS (digital access signaling system) and DPNSS (digital private network signaling system) were developed.

DASS is a signaling system for connecting a PABX to a public network. DPNSS is a digital private network signaling system for setting up PABX networks. Such a network is needed in a multi-site company with thousands or tens of thousands of employees. DPNSS supports a rich set of supplementary services that may be useful in a corporate network. A signaling system that is used to connect a PABX to a public network (DASS or DSS1) does not necessarily need to and usually does not support all the corporate network services.

It is also possible to offer PABX-like corporate network services using a public network switching system for service implementation. This possibility was invented because at some point in the US the Bell Operating Companies providing local telephone services were forbidden to offer PABX services to companies. They tackled the market limitation by implementing similar services in their own Central Offices and they called the service offering: Centrex. Also a combination of PABX and public switch implementation of a private network is possible. In that case the signaling between the concerned PABXs and the public switch is for example DPNSS or Q.SIG. The latter is an international standard based private network signaling system. Q.SIG was defined taking DSS1 as the starting point and extended with features that are typical in a private network.

A PABX is normally connected to an LE using *subscriber criteria*. This means that the LE records the outgoing call minutes or seconds or creates Call Data Records for the calls and the local operator sends a bill for the calls each billing period based on business subscriber tariffs. Outgoing and incoming calls can be processed with the help of the subscriber database in the LE. The set of supplementary services supported across the connection is the set of services that the concerned public network supports. The bill may be charged to the PABX number rather than recorded for each extension of the PABX.

Because DSS1 or any of its variants support sending numbers downstream to a PABX, they all naturally support Direct Dialing In (DDI). DDI means that customers connected to the PBX are directly reachable from the public network by a normal dialing procedure without engaging a human operator on the call. So, with a digital signaling system it is easy to support a setting where in the private network short numbers (e.g. 4 digits) are used for calling inside the private network (private numbering plan) while all extensions still have the DDI capability.

Signaling Efficiency

Signaling efficiency generally refers to the performance aspects of signaling. Different metrics are used for different purposes. If we are interested in the user point of view, *post-dialing delay* is a useful metric. According to E.721 of ITU-T call setup delay (also known as *post-dialing delay* or post-selection delay) is defined as the *interval between entering the last dialed digit and receiving a ringing tone*.

DSS1 and particularly SS7 improved network performance significantly by reducing post-dialing delay. Post dialing delay is a network wide metric: one must always assume some network topology. The relevant requirement is set for the maximum post dialing delay in the network. The ITU-T requirements are the max of 2s for national and 8s for international calls.

Another angle is the *signaling system designers and operators point of view*. In this case we are interested in the use of network transmission capacity and other resources that are needed for signaling. For digital signaling systems a good metric is to look at the size of typical signaling flows created using the signaling system. Flow size can be measured in bits. Bits can be turned into time spent by the signaling flow by a simple formula:

Transfer delay = size in bits/signaling channel speed. (1)

This measure is easy to use for a signaling system. No network topology needs to be assumed. Instead the metric can be used for a one hop signaling connection. For post dialing delay, measurements need to be made across a network or calculations must be based on numerous assumptions about the network. Therefore, post-dialing delay is suitable for characterizing a network rather than a single signaling system.

We can make the following considerations:

All control bits in all calls create signaling traffic – the amount of signaling traffic should be significantly less than the signaling channel capacity

- Users pay for the service: businesswise signaling is pure overhead
- – it follows that the following should apply:

Amount of signaling bits in a call flow << amount of payload bits in the service

- The second reason supporting the above constraint is that otherwise the service delay would grow high.
- Amount of bits in the service is e.g. amount of voice bits in a call.
- If calling is a flat rate service on a broadband network, amount of signaling bits is not that important. Even in this case, the issue of service delay remains relevant.

A DSS1 signaling trace

Let us measure ISDN signaling efficiency for BRI in terms of bits and transfer delay.

The flow has been captured in the TKK Netlab by Vesa Kosonen. Layer two flags and the LAPD checksums are not shown in the captured flow. The presentation uses hexa code instead of binary. In hexa code four bit combinations are replaced by hexadecimal numbers from 0, 1, 2, …A, B, C, D, E and F.

From one ISDN phone (29) to another ISDN phone (32) ENBLOCK SENDING- HEXA representation

```
                    Conn:1 Card:1 Channel:D        52 11:50:09.565
                    RESYNCRONIZATION
                    Conn:1 Card:1 Channel:D        53 11:50:09.567
                    ACTIVATION REQ
                    Conn:1 Card:1 Channel:D        54 11:50:09.570
                    ACTIVATION IND
Conn:1 Card:1 Channel:D        55 11:50:09.580
SABME
                    Conn:1 Card:1 Channel:D        56 11:50:09.601
                    UNNUMBERED ACK
Conn:1 Card:1 Channel:D        57 11:50:09.644
L2:     INFO 00990000080113050403809OA31801836C02008070038033327D0291817E0104
L3:     SETUP 0801130504038090A31801836C02008070038033327D0291817E0104
                    Conn:1 Card:1 Channel:D        58 11:50:09.692
                    RECEIVER READY
                    Conn:1 Card:1 Channel:D        59 11:50:09.817
               L2   INFO :  02990002080193021801 8A
               L3   CALL PROCEEDING :  0801930218018A
Conn:1 Card:1 Channel:D        60 11:50:09.826
RECEIVER READY
                    Conn:1 Card:1 Channel:D        61 11:50:10.370
               L2   INFO :  0299020208019301
               L3   ALERTING :  08019301
Conn:1 Card:1 Channel:D        62 11:50:10.378
RECEIVER READY
                    Conn:1 Card:1 Channel:D        63 11:50:14.019
               L2   INFO 029904020801930729056901071 4247C038090A3
               L3   CONNECT :  0801930729056901071 4247C038090A3
Conn:1 Card:1 Channel:D        64 11:50:14.027
RECEIVER READY
                    Conn:1 Card:1 Channel:D        65 11:50:14.866
               L2   INFO :  0299060208019345080281901E028188
               L3   DISCONNECT :  08019345080281901E028188
Conn:1 Card:1 Channel:D        66 11:50:14.874
RECEIVER READY
Conn:1 Card:1 Channel:D        67 11:50:15.117
L2:      INFO :  009902080801134D08028090
```

L3: **RELEASE** : 0801134D08028090
Conn:1 Card:1 Channel:D 68 11:50:15.166
RECEIVER READY
Conn:1 Card:1 Channel:D 69 11:50:15.282
L2 INFO : 029908040801935A
L3 **RELEASE COMPLETE** : 0801935A
Conn:1 Card:1 Channel:D 70 11:50:15.290
RECEIVER READY
Conn:1 Card:1 Channel:D 71 11:50:20.414
DISCONNECT

By cleaning up we get Table 4.1. LAPD (layer 2 content) is shown in bold. The reminder is Q.931. The used telephone numbers are just two digits. In our comparison, we need to add a reasonable amount of digits to this flow to account for some average number length.

Table 4.1: An ISDN Signaling Flow.

SETUP	**00990000**0801130504038090A31801836C02008070038033327D0291817E0104
CALL PROC	**02990002**0801930218018A
Alerting	**02990202**08019301
Connect	**02990402**080193072905690107142 47C038090A3
Disconnect	**02990602**08019345080281901E028188
Release	**00990208**0801134D08028090
Release Complete	**02990804**0801935A

Counting the bits message by message and accumulating the bits over the whole flow we get Table 4.2.

Table 4.2: Summary of the ISDN flow

A	B	C	D	E	F	G	H
Message	Hexa content	LEN(B)	Bits	N-len =8	Cumu-lative of E	ms of F	Add CK+Deli-miters
SETUP	00990000080113050403809 0A31801836C020080700380 33327D0291817E0104	64	256	352	352	22	24
CALL PROC	029900020801930218018A	22	88	88	440	28	30
Alerting	0299020208019301	16	64	64	504	32	34
Connect	029904020801930729056 90 10714247C038090A3	40	160	160	664	42	44
Discon-nect	0299060208019345080281 9 01E028188	32	128	128	792	50	52
Release	009902080801134D08028090	24	96	96	888	56	58
Release Complete	029908040801935A	16	64	64	952	60	62

The cumulative milliseconds of the ISDN signaling flow, assuming a 16 kbps signaling channel in the 2B+D interface, en-block sending, 8 digits for both the Called Party Number and the Calling Party Number and taking into account both the Q.931 bits and the LAPD bits in the frames that carry the signaling information are in column H of Table 4.2. The result is also presented graphically in Figure 4.4.

Figure 4.4: Summary of an ISDN call setup and Release

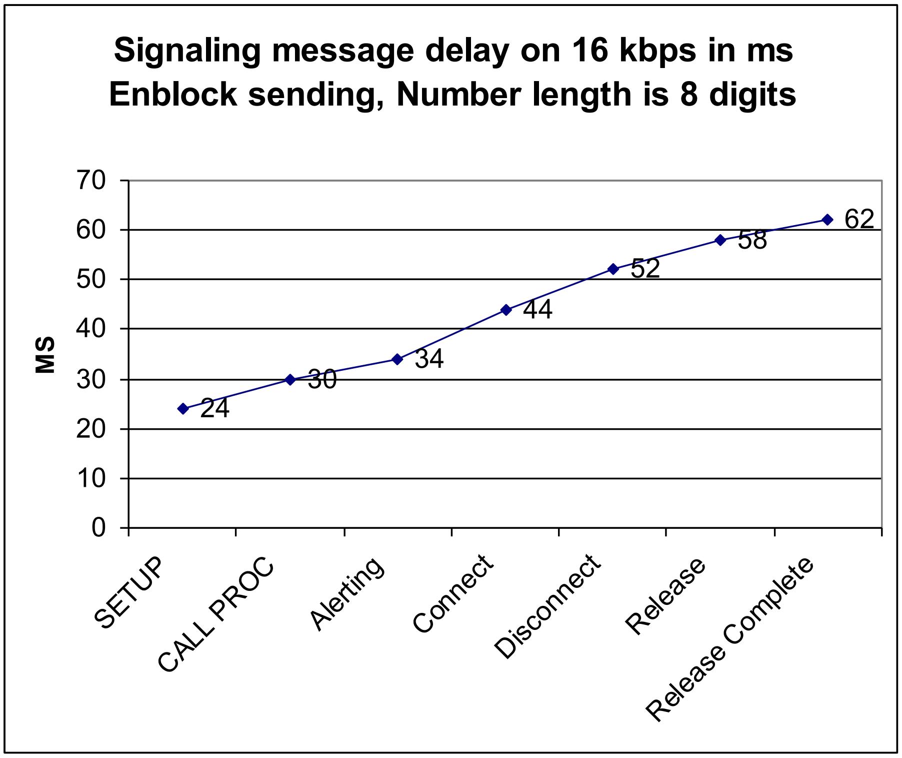

Table 4.4 presents some bit count metrics on the ISDN signaling flow.

ISDN signaling **bitcount** summary		Number length = 8	
	L3	L2+L3	L2 overhead %
Call setup signaling (enblock)	536	664	24 %
Avg message in call setup	134	166	
Call setup and release total	728	952	31 %
Avg message in total	104	136	

Conclusion on ISDN signaling efficiency

The signaling transfer delay occurs due to the need to transfer the signaling bits on a signaling channel of finite speed. We calculate signaling transfer using formula 1.

Transfer delay = size/signaling channel speed.

The total signaling transfer delay for call setup on a 2B+D access is less than 50 ms and total delay including also call release is slightly above 60 ms. Considering that the human reaction time is somewhere between 150ms and 200ms and that people are rarely irritated if they need to wait for less than 2s, ISDN signaling is fast in human terms. When DSS1 used over PRI, signaling delay is about ¼ of the numbers shown here.