Telephony, Modems, and Ethernet

November 8, 2000

1 Telephone Communications

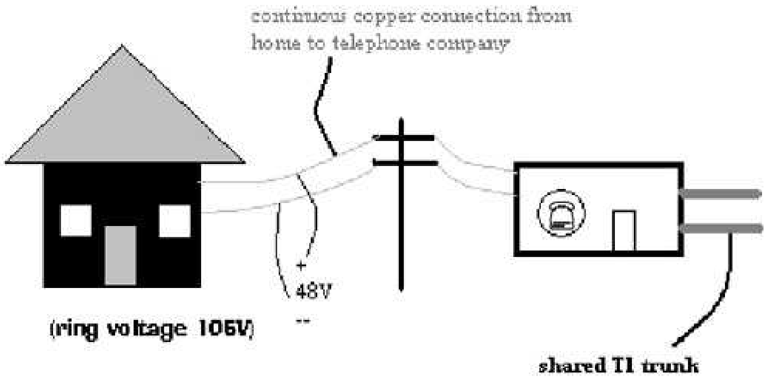
Since most people still dial up to access the Internet, it is important to look at the phone system before studying more complicated Internet protocols.

1.1 Physical Setup

The phone company provides a twisted pair of copper wires from a central office to your house. The phone signal travels across these wires. They also apply a DC voltage of 48V across these lines to power the phone, e.g., to generate tones when you press the buttons on your phone. The DC voltage is supplied by 24V lead-acid cells, thus assuring complete independence from the electric company, which is useful during general power outages.

The phone itself converts the 48V to 5V. The higher transmission voltage is used to minimize power loss. Since there's a fixed power $P = VI$ required by the phone, we have the freedom to choose V such that the current is minimized, thus minimizing $P_{loss} = I^2 R_{line}$.

The actual current drawn is $I \approx 20 - 80$ mA, and the received signal amplitude is $\approx 1V$ (much smaller than the DC component). The amplitude of the transmitted signal is even smaller, $\approx 20mV$.



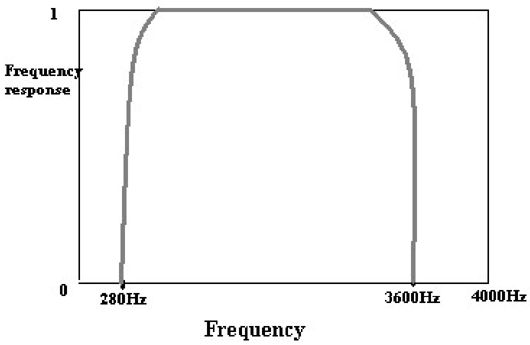
1.2 Telephone "protocol"

When connecting a call, the first thing that needs to be determined is whether it's possible to reserve an end-to-end path from the caller to the number called. If such a path is found, it is reserved and the call will not be interrupted. In rare cases, however, this is not possible and the connection is not established. Thus, the phone system has

- guaranteed bandwidth, and
- a congestion control mechanism ("all circuits are busy").

Once the connection is established, the analog signal travels from the home to the central office where it is sampled and converted to a digital signal. The sampling rate is 8000 samples/sec. This number is a consequence of a sampling theory result indicating that, in order to recover a signal with frequency f, samples must be taken with frequency 2f. (The time between samples is called the Nyquist Interval.) Therefore, the maximum recoverable frequency is 4000Hz. Since the human ear can hear up to 20000Hz, sometimes the effects of the high frequency sounds being clipped are perceptible, i.e., the sound is "tinny".

A low-pass filter is also applied before sampling, and echo suppression and cancellation may be employed as well. This low pass filter cuts off the signal below 280Hz and above 3600Hz. The filter helps to prevent aliasing, a phenomenon in which high-frequency components of a signal at multiples of the sampling frequency amplify some low frequency components when the signal is recovered.



2 Digital data transmission

2.1 Shannon's Limit

Our main concern, however, is data transmission. Shannon derived an upper limit for the quantity of information (number of bits) that can be transmitted over a noisy channel.

$$\frac{I}{T} \leq B \cdot \log_2(1 + \frac{S}{N})$$

where
I = information (bits, a unitless quantity)
T = seconds
B = bandwidth (Hz)
S = power of signal
N = power of noise
As an example, in our phone system, B = 3600Hz - 280Hz ≈3300 Hz. A comprehensive study showed that for at least 50% of residential users, measured S/N ≤35 dB. Thus, S/N ≈35dB = $10^{\frac{35}{10}}$ = 3162. By Shannon's bound, I/T = 3300 Hz $\cdot \log_2$ (3162) ≈ 38,000 bits/sec. Modem designers want to approach this upper limit.

It is important to note here that B, the bandwidth available, is not an upper bound on the rate at which data can be transmitted. If we have a high signal-to-noise ratio, we can get an arbitraririly high throughput, and DSL shows it is practical to do so over the same pair of wires that the plain old telephone system uses. The question is, what keeps the S/N low? The answer has to do with quantization.

2.2 Quantization

When an analog signal is converted to a digital one, there is a finite number of possible values that a sample can take. The error can be as much as half of the quantization step size. For example,a k-bit sample has discrete values corresponding to the integers in the [0, 2^k - 1] interval.

This quantization error is equivalent to a "quantization noise" because even if the channel is noise-free, the signal is not fully recovered. The quantization noise limits the signal to noise ratio.

2.2.1 Why is 35 dB S/N the limit?

The standard rule of thumb is that adding 1 bit of precision to each sample yields an increase of 6 dB in the S/N ration. We'll derive this bound now.

Suppose we are sampling a signal that ranges in voltage between 0 and V_S. Let's first bound the signal strength S. At any instant in time, power is given by

$$P = \frac{v^2}{R} \tag{1}$$

where v is the instantaneous voltage and R is the impedance at the carrier frequency.

We're interested in the average power in the signal. To calculate this precisely, we would have to know something about the shape of the signal. For our upper bound we will just assume that the signal is as strong as it can possibly be all the time, so that the average signal strength is

$$S \leq \frac{V_s^2}{R} \tag{2}$$

Now to compute the noise strength. If the quantization is linear, and there are 2^k possible quantization values, then the distance between two successive values is $\frac{V_S}{2^k - 1}$, which is about $\frac{V_S}{2^k}$. The quantization error will range between $-\frac{V_S}{2*2^k}$ and $+\frac{V_S}{2*2^k}$, and we will assume that the quantization will be distributed uniformly. This is a really big assumption, and the whole argument hangs on it. It's certainly reasonable if the signal is, say, a recording of a piece of music, but it's not so obviously reasonable if the signal is being generated by some sort of $D \rightarrow A$ converter. (Perhaps the converter could be synchronized with the sampler and could always precisely output one of 2^k different voltage values!) But the phone system is peculiar (it filters the signal before sampling it, for example), and after many years of effort the analog signals discovered so far that carry the most information (using quadtrature amplitude modulation, trellis coding, etc.) are horrendously complicated, and the quantization error really looks random and uniform. So the average strength of the noise is given by the integral

$$N = \int_{-V_s/(2*2^k)}^{+V_s/(2*2^k)} \frac{\frac{v^2}{R}}{\frac{V_s}{2*2^k}} dv \tag{3}$$

(Here I've assumed the same impedance R for both the signal and the noise.) This is a "mean square" calculation. It solves to

$$N = \frac{V_s^2}{12 * 2^{2k}} \tag{4}$$

Now

$$\frac{S}{N} = 12 * 2^{2k} = 12 * 4^k \tag{5}$$

In other words, the $\frac{S}{N}$ ratio increases by a factor of 4 for each bit of precision in the samples. Converting to dB, we have

$$10 \log_{10} \frac{S}{N} \leq 10 \log_{10} 12 * 4^k = 10.8 + 6.02k \tag{6}$$

This is where the 6dB/bit figure came from. The additive 10.8 is an artifact of our assumption that the signal is always at full strength. Empirically, the number that is used by modem designers is more like 1.17, rather than 10.8.

Finally, going back to Shannon's Limit, we have

$$\frac{I}{T} \leq B \log_2 (1 + \frac{S}{N}) \tag{7}$$

Let's drop the "1" in "$1 + \frac{S}{N}$" as "$\frac{S}{N}$" is large. Plugging in $B = 3200Hz$, and $\frac{S}{N} = 12 * 4^k = 12 * 4^8$, we have

$$\frac{I}{T} \leq 3200 * 2k + \log_2 12 = 3200 * 19.6 = 62700 \text{bits/sec} \tag{8}$$

So this gives us an upper bound on the data rate based solely on quantization error (and the assumption that quantization is random and uniform). In reality, there are are other sources of noise and distortion that bring this upper bound down to approximately 36000 bits/sec. For example, 3200Hz is a very optimistic estimate of the available bandwidth. The low pass filter is not perfectly sharp, and actually significantly reduces the signal outside a range of about 2900Hz. Furthermore, we didn't take into account the fact that quantization of a phone signal is non-linear (described later), and other sources of distortion.

2.3 T1 Trunk

Once the signal reaches the central office, it is combined with other signals that are to be carried over a T1 trunk (two twisted pairs of copper wires). The samples from each signal are interleaved and packed into "frames". The phone system reserves a fixed position within each frame for each call; the sample placed in this position is the only information about the call that is carried in the frame, which a small exception described below. A T1 trunk supports 24 voice grade lines simultaneously. Hence each frame consists of 24 8-bit samples. At the end of each frame is a parity bit, so the total number of bits in a frame is 193. The T1 trunk transmits 8000 frames/second, for a total speed of 1.544 Mbits/sec

8 8 ... 8 1

A frame : (24 * 8) + 1 = 193 bits (1 8-bit sample from each of 24 voice grade lines)

Every 6th frame, each sample loses its least significant bit. This bit is "stolen" to indicate whether the call is still in progress. The effect is a small amount of noise in the reproduced signal that is not discernible by a human listener. However, the loss of this bit makes packing bytes of data into frames inconvenient, since every sixth one loses a bit. This (and other reasons discussed below) is why the current limit for modem speeds is 56Kbps and not 64Kbps – data is sent 7 bits at-a-time rather than 8.

3 Modems

A modem accepts a sequence of signals and modulates into analog signals to be sent over the voice channel. The "mo-dem" has two main functions:

- modulate : convert from digital to analog
- demodulate: convert from analog to digital

Modem function also includes compression and error detection and correction. Current compress ratios are about 2:1 for text and binary executables, using a form of Ziv-Lempel. The stated speed of a modem does not include this compression factor.

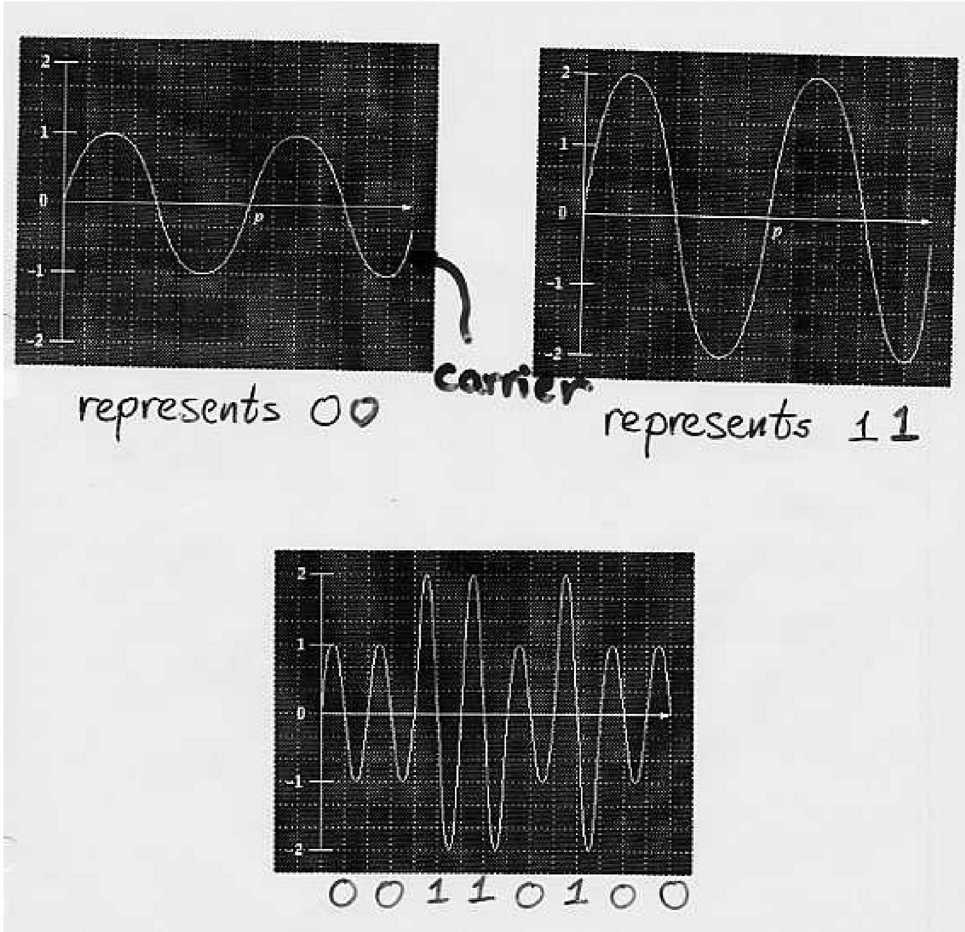
Three different techniques are used for modulation:

1. Amplitude modulation
2. Phase shift keying
3. Trellis coding

The goal is to encode 9 bits per symbol and to transmit appoximately 4000 symbols per second, thus approaching the upper bound for data transmission given by Shannon's limit.

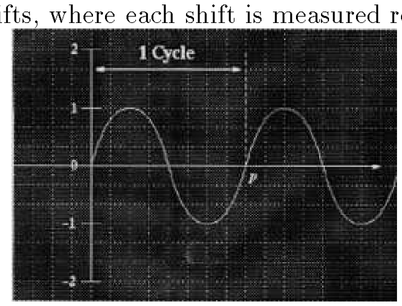
3.1 Amplitude modulation

The idea behind amplitude modulation is simple: In each full wave on the carrier frequency, a 0 is encoded by a magnitude of 1, and a 1 is encoded by doubling the magnitude. The message can be read off the resulting signal as follows:

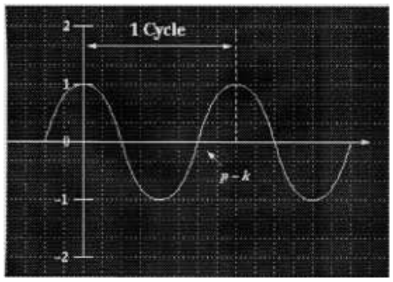


3.2 Phase-Shift Keying

Instead of changing amplitude, phase-shift keying relies on changing the phase of the carrier wave, i.e., shifting it in time. Pairs of bits are encoded by consecutive shifts, where each shift is measured relative to the phase of the previous wave.



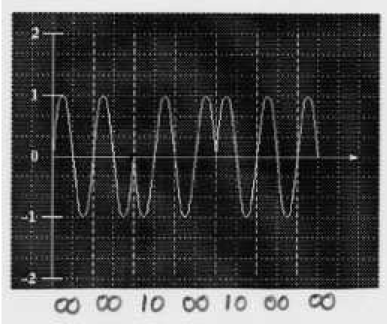
00 : no shift



1/4 shift left - represents 01

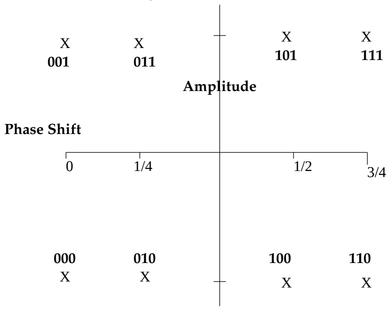
shift	*value*
0	00
1/4	01
1/2	10
3/4	11

The shift is relative to the previous value.

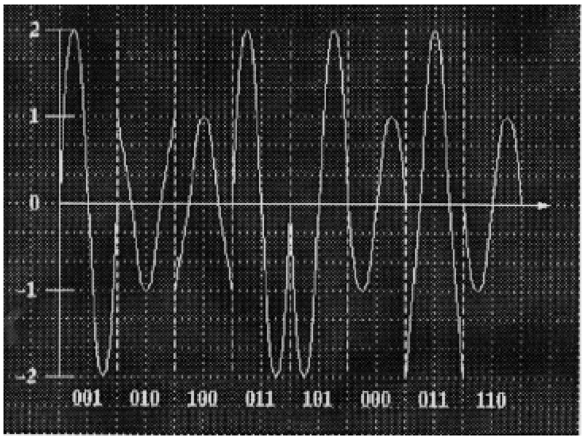


3.3 Quadrature amplitude modulation

Used together, phase-shift keying and amplitude modulation are called quadrature amplitude modulation. In our example, the two techniques can be combined to encode 3 bits/symbol, as follows:



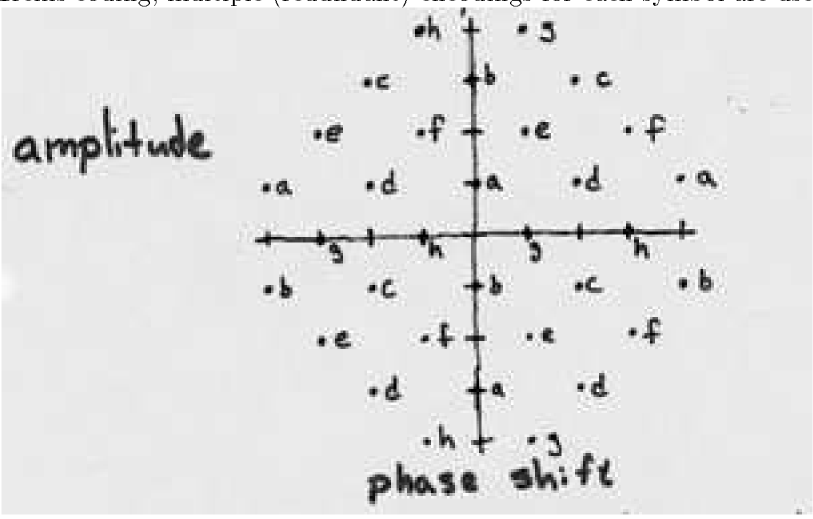
Quadrature amplitude modulation



3200 baud (symbols/sec)* 9 bits/symbol = 28,800 bits/sec
3429 symbols/sec * 9 bits/symbol = 30861 bits/sec

3.4 Trellis coding

In Trellis coding, multiple (redundant) encodings for each symbol are used.

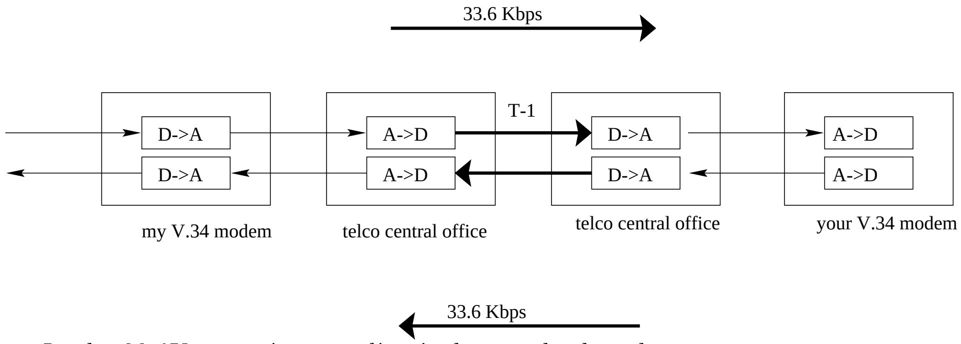


Consecutive symbols in the data stream are mapped to well-separated points in the plane. As we're encoding, we look at the previous symbols and see where we can map the next one. The non-linear improvement in symbol recognition compensates for inefficiency of redundant encoding .

3.5 Do 56K modems violate Shannon's Limit?

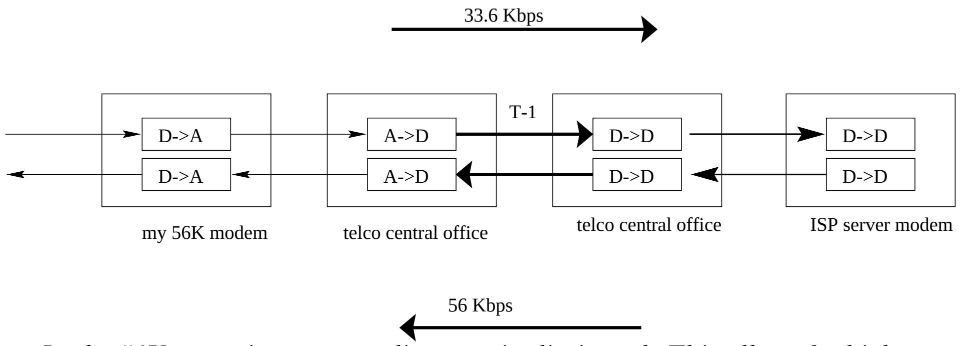
No! Analog-to-digital conversion is avoided altogether. There is no sampling. This is possible because the Internet Service Provider has a direct digital line to the telephone company:

3.5.1 33.6 K Scenario



In the 33.6K scenario, sampling is done at both ends.

3.5.2 56K Scenario



In the 56K scenario, one sampling step is eliminated. This allows for higher download speeds, but the upload speed remains 33.6 Kbs

3.6 Why 56 and not 64 K?

Theoretically, we could transmit 8000 frames/sec $\cdot 8$ bits/frame $= 64000$ bits/sec. However, we've already mentioned the inconvenient loss of one bit from every sixth sample. In addition to this, the telephone system is optimized for voice communication. In the sampling step, the A->D conversion performed at the central office is non-linear; instead, it follows the function below. In the high end of the signal a large change in voltage results in a small change in the sample value. This provides better reconstruction at lower amplitudes (people usually speak quietly, and don't mind a little distortion when the speaker is yelling).

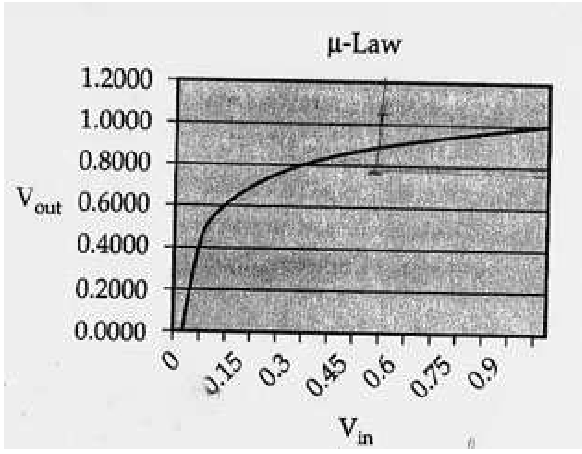
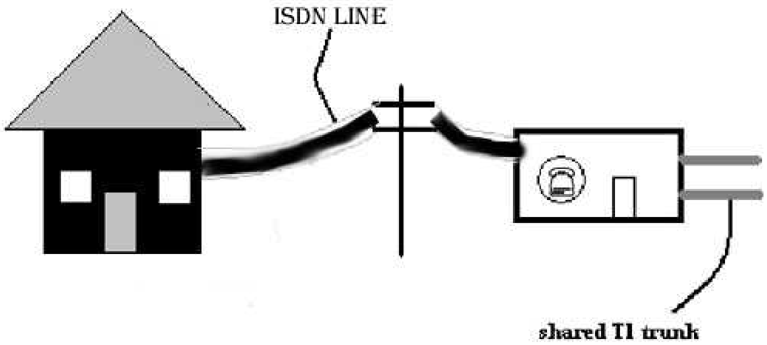


Figure : Non-linear A->D conversion.

The final D->A conversion function is the inverse of the one shown above. This is appropriate for voice but becomes a problem because digital data becomes harder to decode, as the voltage levels of some of the 256 different 8 - bit values are too close together. The solution is to only use the best 128 values (7 bits) , for a total bit-rate of $8000 \cdot 7 = 56000$.

4 ISDN (Integrated Services Digital Network)



The digital signal is sent directly, encoded using 4 voltage levels. There are three channels: 2b + d, where

d = demand 16000 bits/sec

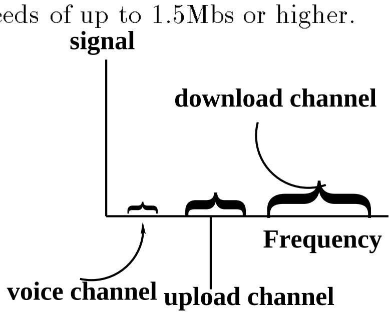
b = bearer (of data) : 64 000 bits/sec, for an aggregate bandwidth 144000 bits/sec.

4.1 DSL

Fro the Digital Subscriber Line, the data rates are much higher than the regular voice channel.The frequency band is divided into three channels: one for voice, one for upload and one for download. DSL also works over one twisted pair of copper wires.

4.1.1 ADSL (Asynchronous Digital Subscriber Line)

Since download needs are usually higher than upload needs, ADSL dedicates a larger portion of the frequency band to the download channel, allowing download speeds of up to 1.5Mbs or higher.

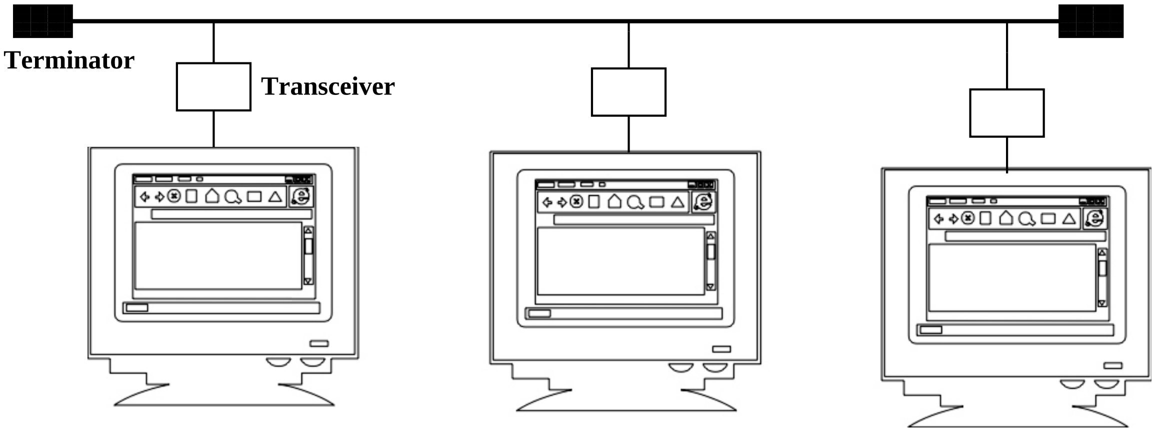


4.1.2 HDSL (High-bit-rate Digital Subscriber Line)

HDSL divides the frequency band into two equal channels for upload and download.

5 Ethernet (XEROX PARC)

The Ethernet originated in XEROX PARC; it was originally based on a 50 ohm coaxial cable used as a bus to which many transceivers can connect.



The terminator at the end of the cable simulates an infinitely long cable. (Without terminators, signals bounce back from the ends of the cable.)

Ethernet speeds:

10 M bits/sec – 1980

100 Mbits/sec to 1Gbits/sec – today

10 base 5, 2 = coax;

10 base T = twisted pair

10 base F = fiber

Everybody listens to the same bus and uses the same bus to transmit the data. A protocol is needed for deciding when to transmit a packet.

5.1 The CSMA/CD protocol

(Carrier Sense Multiple Access with Collision Detect)

To transmit a packet:

1. Wait until media is quiet (9.6 μs pause between packets). (why not wait a random amount of time ? Ethernet is usually quiet so collisions are unlikely.)
2. Begin transmission
3. Listen for collision while transmitting

On collision

1. Send collision notice (a special, out-of-bound signal that can be recognized by everybody)
2. Wait for a random amount of time; retransmit
3. Double waiting time with each collision (binary exponential backoff)
4. Give up after 16 collisions – best effort delivery

No priority is given to the sender of previous packet; the protocol is designed to prevent starvation.

5.2 Ethernet Frame (Packet) Structure

56 bits	8 bits	48 bits	48 bits	16 bits	46-1500 bytes	32 bits
preamble	start frame delimiter	destination	source	type/length	data ...... {pad}	CRC

The preamble (56 bits) allows the listener and the sender to synchronize. The destination and the source address are the hardware addresses. The fifth field (16 bits) is used for the type information in the Ethernet II standard, and for length in the IEEE 802.3 standard. The pad is necesary because the minimum frame size is 64 bytes. The 32 bits at the end of the frame are the cyclic redundancy check bits.

Every Ethernet transceiver has a built in 48-bit hardware address (e.g., 0C:1B:99:AC:00:11). The first 24 bits specify the manufacturer. There are over 40 million installed. There is some concern that we will run out of distinct ethernet hardware addresses. As we shall see, however, the shortage of IP addresses is much more urgent.