



Data Communication

#6 Data Rate Limit and Performance

Susmini I. Lestaringati, M.T

Data Rate Limits

- A very important consideration in data communications is how fast we can send data, in bits per second, over a channel. Data rate depends on three factors:
 1. The bandwidth available
 2. The level of the signals we use
 3. The quality of the channel (the level of noise)
- Two theoretical formulas were developed to calculate the data rate:
one by Nyquist for a noiseless channel, another by Shannon for a noisy channel.

Noiseless Channel : Nyquist Bit Rate

- For a noiseless channel, the Nyquist bit rate formula defines the theoretical maximum bit rate

$$\text{Bit Rate} = 2 \times B \times \log_2 L$$

- In this formula, B is the bandwidth of the channel, L is the number of signal levels used to represent data, and Bit Rate is the bit rate in bits per second.
- According to the formula, we might think that, given a specific bandwidth, we can have any bit rate we want by increasing the number of signal levels. Although the idea is theoretically correct, practically there is a limit.
 - When we increase the number of signal levels, we impose a burden on the receiver. If the number of levels in a signal is just 2, the receiver can easily distinguish between a 0 and a 1. If the level of a signal is 64, the receiver must be very sophisticated to distinguish between 64 different levels. In other words, increasing the levels of a signal reduces the reliability of the system.

Increasing the levels of a signal may reduce the reliability of the system.

Example

1. Consider a noiseless channel with a bandwidth of 3000 Hz transmitting a signal with two signal levels. The maximum bit rate can be calculated as

$$\text{BitRate} = 2 \times 3000 \times \log_2 2 = 6000 \text{ bps}$$

2. Consider the same noiseless channel transmitting a signal with four signal levels (for each level, we send 2 bits). The maximum bit rate can be calculated as

$$\text{BitRate} = 2 \times 3000 \times \log_2 4 = 12,000 \text{ bps}$$

Example

We need to send 265 kbps over a noiseless channel with a bandwidth of 20 kHz. How many signal levels do we need?

Solution

We can use the Nyquist formula as shown:

$$265,000 = 2 \times 20,000 \times \log_2 L$$
$$\log_2 L = 6.625 \quad L = 2^{6.625} = 98.7 \text{ levels}$$

Since this result is not a power of 2, we need to either increase the number of levels or reduce the bit rate. If we have 128 levels, the bit rate is 280 kbps. If we have 64 levels, the bit rate is 240 kbps.

Noisy Channel : Shannon Capacity

- In reality, we cannot have a noiseless channel; the channel is always noisy.
- In 1944, Claude Shannon introduced a formula, called the Shannon capacity, to determine the theoretical highest data rate for a noisy channel:

$$\text{Capacity} = \text{Bandwidth} \times \log_2 (1 + \text{SNR})$$

- In this formula, bandwidth is the bandwidth of the channel, SNR is the signal-to-noise ratio, and capacity is the capacity of the channel in bits per second.
- Note that in the Shannon formula there is no indication of the signal level, which means that no matter how many levels we have, we cannot achieve a data rate higher than the capacity of the channel. In other words, the formula defines a characteristic of the channel, not the method of transmission.

Example

Consider an extremely noisy channel in which the value of the signal-to-noise ratio is almost zero. In other words, the noise is so strong that the signal is faint.

For this channel the capacity C is calculated as

$$C = B \log_2 (1 + \text{SNR}) = B \log_2 (1 + 0) = B \log_2 1 = B \times 0 = 0$$

This means that the capacity of this channel is zero regardless of the bandwidth. In other words, we cannot receive any data through this channel.

Example

- We can calculate the theoretical highest bit rate of a regular telephone line. A telephone line normally has a bandwidth of 3000 Hz (300 to 3300 Hz) assigned for data communications. The signal-to-noise ratio is usually 3162. For this channel the capacity is calculated as

$$\begin{aligned}C &= B \log_2 (1 + \text{SNR}) \\&= 3000 \log_2 (1 + 3162) = 3000 \log_2 3163 \\&= 3000 \times 11.62 = 34,860 \text{ bps}\end{aligned}$$

This means that the highest bit rate for a telephone line is 34.860 kbps. If we want to send data faster than this, we can either increase the bandwidth of the line or improve the signal-to-noise ratio.

Example

- The signal-to-noise ratio is often given in decibels. Assume that $\text{SNR}_{\text{dB}} = 36$ and the channel bandwidth is 2 MHz. The theoretical channel capacity can be calculated as

$$\begin{aligned}\text{SNR}_{\text{dB}} &= 10 \log_{10} \text{SNR} \quad \rightarrow \quad \text{SNR} = 10^{\text{SNR}_{\text{dB}}/10} \quad \rightarrow \quad \text{SNR} = 10^{3.6} = 3981 \\ C &= B \log_2 (1 + \text{SNR}) = 2 \times 10^6 \times \log_2 3982 = 24 \text{ Mbps}\end{aligned}$$

Example

- For practical purposes, when the SNR is very high, we can assume that $\text{SNR} + 1$ is almost the same as SNR. In these cases, the theoretical channel capacity can be simplified to

$$C = B \times \frac{\text{SNR}_{\text{dB}}}{3}$$

- For example, we can calculate the theoretical capacity of the previous example as

$$C = 2 \text{ MHz} \times \frac{36}{3} = 24 \text{ Mbps}$$

Example

- We have a channel with a 1-MHz bandwidth. The SNR for this channel is 63. What are the appropriate bit rate and signal level?
- Solution

First, we use the Shannon formula to find the upper limit.

$$C = B \log_2 (1 + \text{SNR}) = 10^6 \log_2 (1 + 63) = 10^6 \log_2 64 = 6 \text{ Mbps}$$

The Shannon formula gives us 6 Mbps, the upper limit. For better performance we choose something lower, 4 Mbps, for example. Then we use the Nyquist formula to find the number of signal levels.

$$4 \text{ Mbps} = 2 \times 1 \text{ MHz} \times \log_2 L \quad \rightarrow \quad L = 4$$

**The Shannon capacity gives us the upper limit;
the Nyquist formula tells us how many signal levels we need.**

Performance

- One important issue in networking is the performance of the network—how good is it? We discuss quality of service, an overall measurement of network performance, in greater detail in Chapter 24. In this section, we introduce terms that we need for future chapters.
- Topics discussed in this section:
 - **Bandwidth**
 - **Throughput**
 - **Latency (Delay)**
 - **Bandwidth-Delay Product**

Bandwidth

- One characteristic that measures network performance is bandwidth. However, the term can be used in two different contexts with two different measuring values: bandwidth in hertz and bandwidth in bits per second.

- **Bandwidth in Hertz**

We have discussed this concept. Bandwidth in hertz is the range of frequencies contained in a composite signal or the range of frequencies a channel can pass. For example, we can say the bandwidth of a subscriber telephone line is 4 kHz.

- **Bandwidth in Bits per Seconds**

The term bandwidth can also refer to the number of bits per second that a channel, a link, or even a network can transmit. For example, one can say the bandwidth of a Fast Ethernet network (or the links in this network) is a maximum of 100 Mbps. This means that this network can send 100 Mbps.

- **Relationship**

There is an explicit relationship between the bandwidth in hertz and bandwidth in bits per seconds. Basically, an increase in bandwidth in hertz means an increase in bandwidth in bits per second. The relationship depends on whether we have baseband transmission or transmission with modulation. We discuss this relationship in Chapters 4 and 5.

- In networking, we use the term bandwidth in two contexts.
 - The first, bandwidth in hertz, refers to the range of frequencies in a composite signal or the range of frequencies that a channel can pass.
 - The second, bandwidth in bits per second, refers to the speed of bit transmission in a channel or link.

- **Example**
 - The bandwidth of a subscriber line is 4 kHz for voice or data. The bandwidth of this line for data transmission can be up to 56,000 bps using a sophisticated modem to change the digital signal to analog.
 - If the telephone company improves the quality of the line and increases the bandwidth to 8 kHz, we can send 112,000 bps by using the same technology.

Throughput

- The **throughput** is a measure of how fast we can actually send data through a network. Although, at first glance, bandwidth in bits per second and throughput seem the same, they are different.
- A link may have a bandwidth of B bps, but we can only send T bps through this link with T always less than B .
- In other words, the **bandwidth** is a potential measurement of a link; the **throughput** is an actual measurement of how fast we can send data.
- For example, we may have a link with a bandwidth of 1 Mbps, but the devices connected to the end of the link may handle only 200 kbps. This means that we cannot send more than 200 kbps through this link.

Example

- A network with bandwidth of 10 Mbps can pass only an average of 12,000 frames per minute with each frame carrying an average of 10,000 bits. What is the throughput of this network?

- Solution

We can calculate the throughput as

$$\text{Throughput} = \frac{12,000 \times 10,000}{60} = 2 \text{ Mbps}$$

The throughput is almost one-fifth of the bandwidth in this case.

Latency (Delay)

- The **latency** or **delay** defines how long it takes for an entire message to completely arrive at the destination from the time the first bit is sent out from the source.
- We can say that latency is made of four components: *propagation time, transmission time, queuing time and processing delay.*

**Latency = propagation time + transmission time +
queuing time + processing delay**

- **Propagation Time**

- Propagation time measures the time required for a bit to travel from the source to the destination. The propagation time is calculated by dividing the distance by the propagation speed.

$$\text{Propagation Time} = \frac{\text{Distance}}{\text{Propagation Speed}}$$

- The propagation speed of electromagnetic signals depends on the medium and on the frequency of the signal. For example, in a vacuum, light is propagated with a speed of 3×10^8 m/s. It is lower in air; it is much lower in cable.

- **Transmission Time**

- In data communications we don't send just 1 bit, we send a message. The first bit may take a time equal to the propagation time to reach its destination; the last bit also may take the same amount of time. However, there is a time between the first bit leaving the sender and the last bit arriving at the receiver. The first bit leaves earlier and arrives earlier; the last bit leaves later and arrives later. The time required for transmission of a message depends on the size of the message and the bandwidth of the channel.

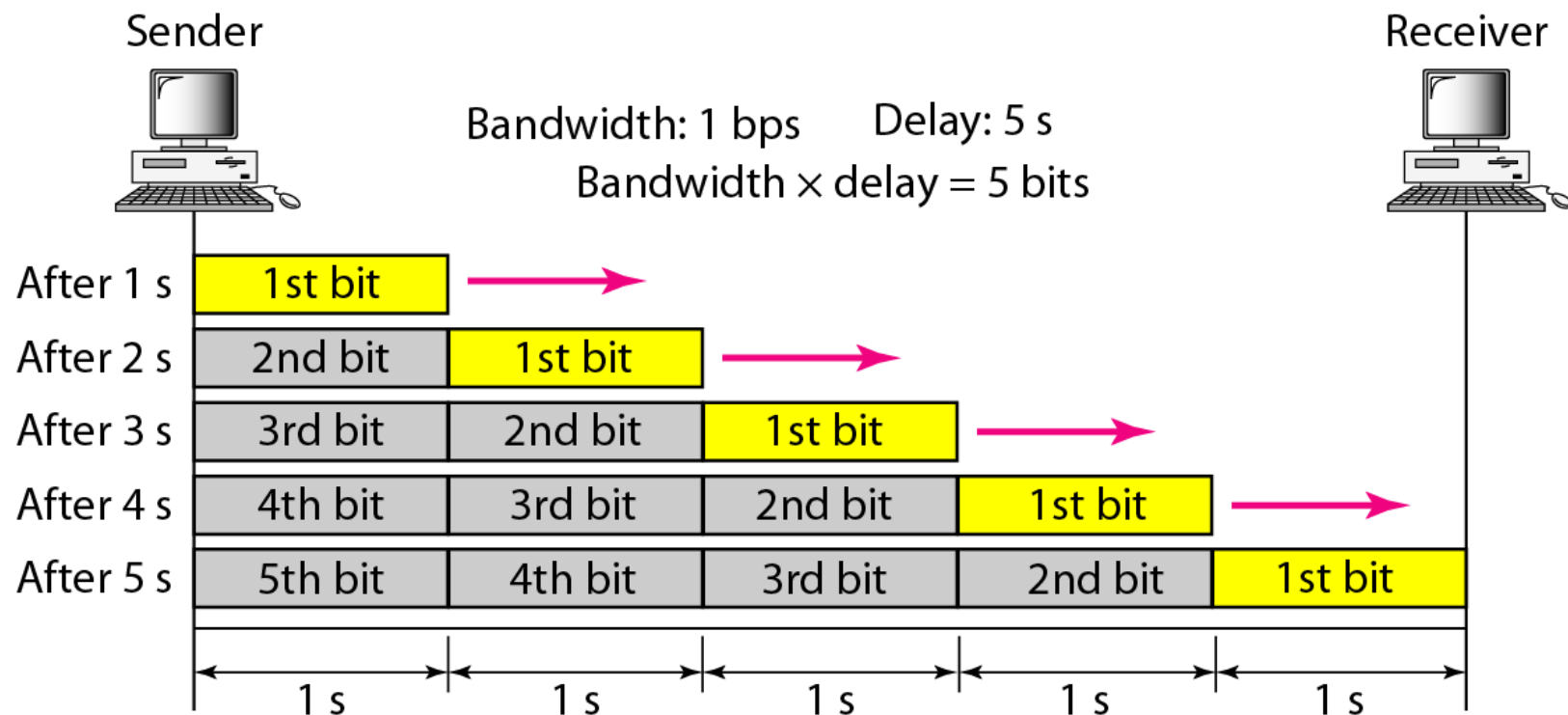
$$\text{Transmission Time} = \frac{\text{Message Size}}{\text{Bandwidth}}$$

- **Queuing Time**

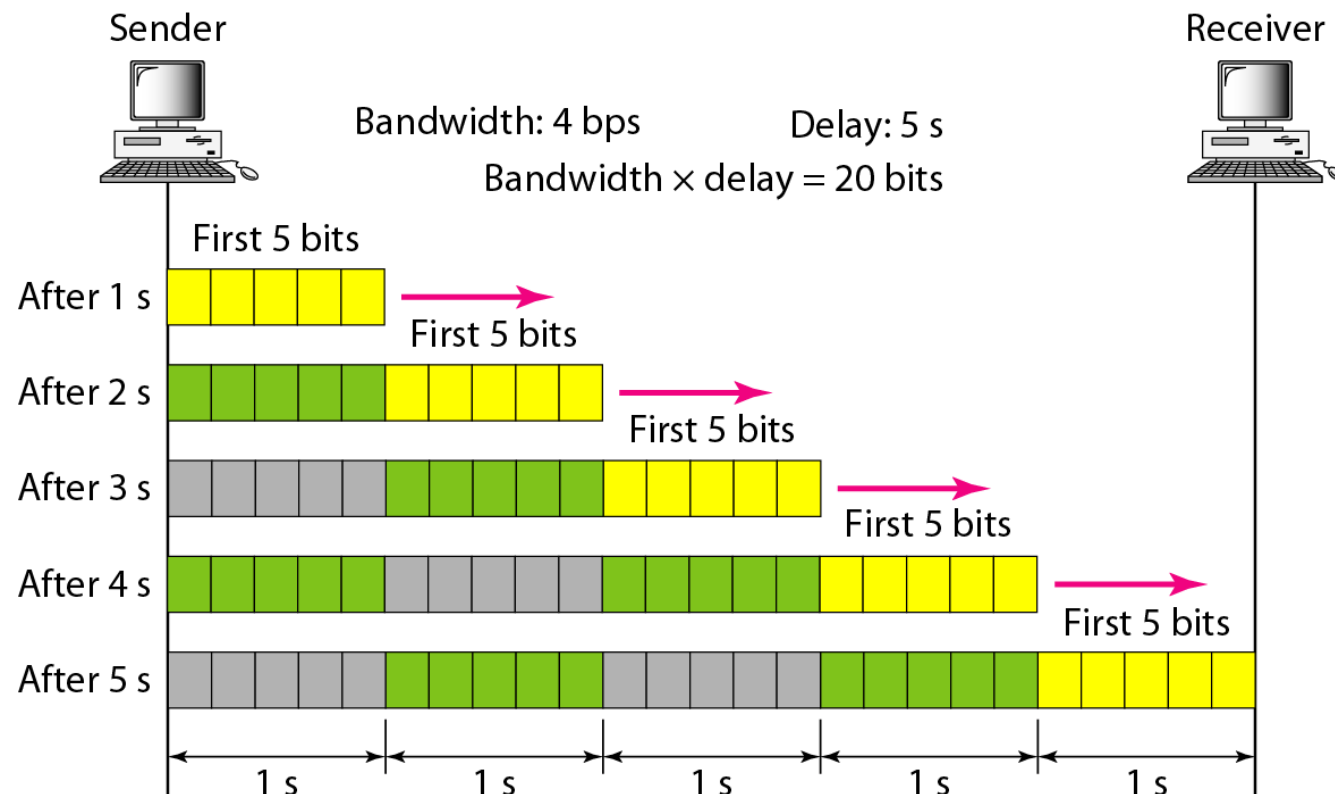
- The third component in latency is the queuing time, the time needed for each intermediate or end device to hold the message before it can be processed.
- The queuing time is not a fixed factor; it changes with the load imposed on the network. When there is heavy traffic on the network, the queuing time increases. An intermediate device, such as a router, queues the arrived messages and processes them one by one. If there are many messages, each message will have to wait.

Bandwidth Delay Product

- Bandwidth and delay are two performance metrics of a link. However, as we will see in this chapter and future chapters, what is very important in data communications is the product of the two, the bandwidth-delay product
- Filling the link with bits for case 1



- Now assume we have a bandwidth of 4 bps. Figure 3.32 shows that there can be maximum $4 \times 5 = 20$ bits on the line. The reason is that, at each second, there are 4 bits on the line; the duration of each bit is 0.25 s.
- Filling the link with bits in case 2

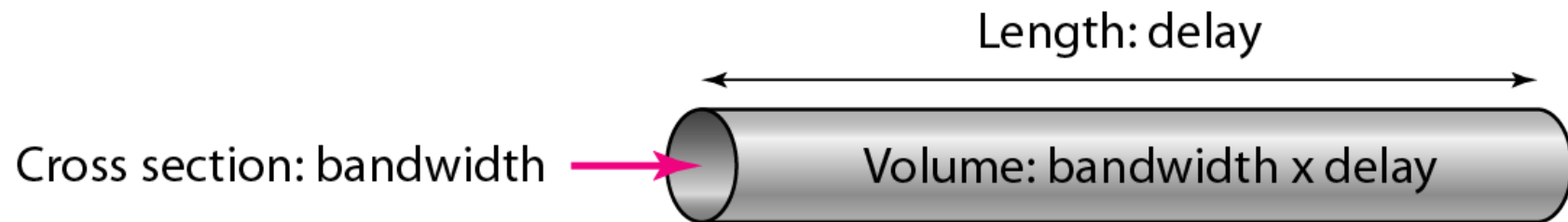


- The above two cases show that the product of bandwidth and delay is the number of bits that can fill the link. This measurement is important if we need to send data in bursts and wait for the acknowledgment of each burst before sending the next one.
- To use the maximum capability of the link, we need to make the size of our burst 2 times the product of bandwidth and delay; we need to fill up the full-duplex channel (two directions). The sender should send a burst of data of $(2 \times \text{bandwidth} \times \text{delay})$ bits. The sender then waits for receiver acknowledgment for part of the burst before sending another burst. The amount $2 \times \text{bandwidth} \times \text{delay}$ is the number of bits that can be in transition at any time.

The bandwidth-delay product defines the number of bits that can fill the link.

Concept of Bandwidth Delay Product

- We can think about the link between two points as a pipe. The cross section of the pipe represents the bandwidth, and the length of the pipe represents the delay. We can say the volume of the pipe defines the bandwidth-delay product, as shown in Figure



Jitter

- Another performance issue that is related to delay is jitter.
- We can roughly say that jitter is a problem if different packets of data encounter different delays and the application using the data at the receiver site is time-sensitive (audio and video data, for example).
- If the delay for the first packet is 20 ms, for the second is 45 ms, and for the third is 40 ms, then the real-time application that uses the packets endures jitter. We discuss jitter in greater detail in Chapter 29.