

Latency on a Switched Ethernet Network

1 Introduction

This document serves to explain the sources of latency on a switched Ethernet network and describe how to calculate cumulative latency as well as provide some real world examples. Latency in a communications network is defined as the time it takes for a message to traverse the network from the transmitter to the receiver. In certain applications, like voice or real-time automation, the network must guarantee a certain maximum latency or the application may not work in a satisfactory manner or, worse, may fail outright. Switched Ethernet networks have several sources of latency: 1) store and forward, 2) switch fabric processing, 3) wireline transmission, and 4) frame queuing. All of these latencies except for queuing are deterministic and yet the effects of frame queuing can also be calculated providing one knows the nature of all sources of traffic on the network.

2 Sources of Latency

2.1 Store and Forward Latency (L_{SF})

Store and forward refers to the basic operating principle of an Ethernet switch. The term is descriptive of its actual operation: the switch stores the received data in memory until the entire frame is received. The switch then transmits the data frame out the appropriate port(s). The latency this introduces is proportional to the size of the frame being transmitted and inversely proportional to the bit rate as follows:

$$L_{SF} = FS / BR$$

where L_{SF} is the store and forward latency, FS is the frame size in bits, and BR is the bit rate in bits/s. For the maximum size Ethernet frame (1500 bytes) at 100 Mbps the latency is 120 μ s. For comparison, the minimum size frame (64 bytes) at Gigabit speeds has a latency of just 0.5 μ s.

2.2 Switch Fabric Latency (L_{SW})

The internals of an Ethernet switch are known as the *switch fabric*. The switch fabric consists of sophisticated silicon that implements the store and forward engine, MAC address table, VLAN, and CoS, among other functions. The fabric introduces delay when executing the logic that implements these functions. The switch fabric latency on RuggedSwitch® products is 5.2 μ s.

2.3 Wireline Latency (L_{WL})

Bits transmitted along a fiber optic link travel at about $\frac{2}{3}$ of the speed of light (3×10^8 m/s) . When very long distance Ethernet links are deployed, this delay can become significant. The one way latency for a 100km link works out to:

$$L_{WL} = 1 \times 10^5 \text{ m} / (0.67 \times 3 \times 10^8 \text{ m/s}) \approx 500 \text{ } \mu\text{s}$$

Note that for the distances involved in local area networks, this delay becomes trivial compared with the other contributions to latency.

Latency on a Switched Ethernet Network

2.4 Queuing Latency (L_Q)

Ethernet switches use queues in conjunction with the store and forward mechanism to eliminate the problem of frame collisions that used to exist on broadcast Ethernet networks. Queuing introduces a non-deterministic factor to latency since it can often be very difficult to predict exact traffic patterns on a network. *Class of Service* (CoS) introduces a priority scheme to Ethernet frames to help mitigate queuing latency. It is a best-effort service, however, and cannot guarantee quality of service, since multiple frames at the highest priority level must still be queued relative to one another. Another consideration is that if a lower priority frame has already started transmission, then that frame must be completed before the switch may begin transmitting the higher priority frame.

Calculating with absolute certainty the worst case latency for any Ethernet frame can be challenging. It requires detailed knowledge about all sources of traffic on the network. Specifically, one must know the maximum frame size transmitted by any device, the CoS priority of frames, and the time distribution and rate of frames. In an arbitrary communications network, little of this information is known and some assumptions have to be made.

Estimating the average latency for an Ethernet frame is simple. For a network with no traffic load, the queuing latency for a frame will be nil. For a loaded network, one can assume that the likelihood of a frame already in the queue is proportional to the network load. The average queuing latency can then be estimated as:

$$L_Q = (\text{Network Load}) \times L_{SF(\max)}$$

where L_Q is the average latency due to queuing, *Network Load* is the fractional load relative to full network capacity and $L_{SF(\max)}$ is the store and forward latency of a full-size (1500 byte) frame. For example, a network with 25% load would have an average queuing latency of:

$$L_Q = 0.25 \times [12000 \text{ bits} / 100\text{Mbps}] = 30\mu\text{s}$$

2.5 Total Worst-Case Latency Calculation (L_{TOTAL})

The latency sources described above are duplicated for every switch that an Ethernet frame must traverse on its journey from source to destination. Hence the general calculation for worst-case latency in a switched Ethernet network is expressed as:

$$L_{TOTAL} = [L_{SF} + L_{SW} + L_{WL} + L_Q] \times N_{\text{switches}}, \text{ or } L_{TOTAL} = \sum_{\text{switches}} [L_{SF} + L_{SW} + L_{WL} + L_Q],$$

where each contribution to latency is considered separately for each switch in the path. The calculation may be simplified considerably if one considers the case where only one traffic source has a high priority and is infrequent enough that multiple frames of that type need never be queued at any switch in the network. In this situation, the worst-case queuing latency is exactly that due to one maximum sized frame in each switch in the path. The worst-case latency then simplifies to:

$$L_{TOTAL} = [(FS / BR) + L_{SW} + (FS_{MAX} / BR)] \times N_{\text{switches}} + L_{WL(\text{total})}$$

where FS is the size of the high-priority frame being considered, in bits, and $L_{WL(\text{total})}$ is the latency due to the cumulative wireline distance from transmitter to receiver.

Latency on a Switched Ethernet Network

3 Examples

3.1 Simple Ethernet Network

Consider a simple network consisting of several nodes connected to one switch, all at 100Mbps.

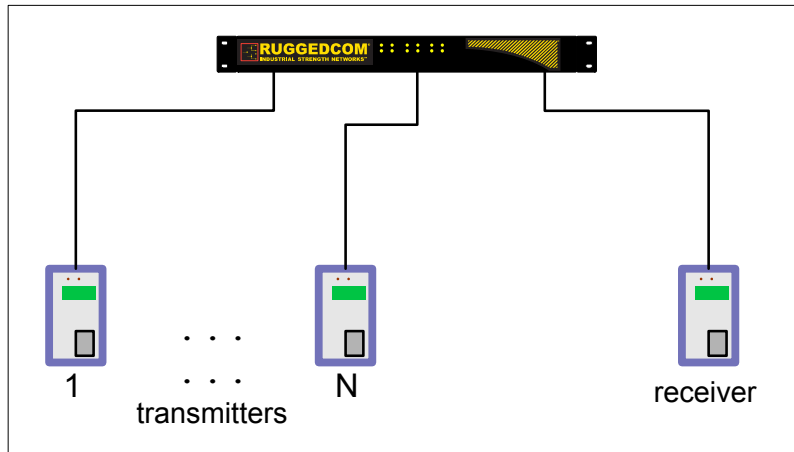


Figure 1: Generic Ethernet Example

Under light network load, one need not consider queuing effects other than for a single frame. The latencies of a minimum-size frame of 64 bytes and a full-size frame of 1518 bytes would be delayed by, respectively:

$$L (64 \text{ byte frame}) = [576 \text{ bits} / 100\text{Mbps}] + 5.2\mu\text{s} = 11\mu\text{s}$$

$$L (1518 \text{ byte frame}) = [12208 \text{ bits} / 100\text{Mbps}] + 5.2\mu\text{s} = 127\mu\text{s}$$

Under heavy load it might become necessary for the switch to queue frames in buffer memory, introducing additional latency. For example, if at the same time every one of N nodes were to transmit a full-size frame to one receiving node on the same switch, then N frames would be in the queue to be transmitted. The worst-case latency, the time from the start of reception of the last frame to be received by the switch to the start of its transmission, is simplified as:

$$\begin{aligned} L (\text{worst-case}) &= (FS \times N / BR) + L_{SW} \\ &= (12304 \text{ bits} \times 15 / 100\text{Mbps}) + 5.2\mu\text{s} = 1.85\text{ms} \end{aligned}$$

Note that here, the 96 bit inter-frame gap is included in FS. In the case of a 16-port switch, N=15. The best-case latency would be 127µs, as previously.

Latency on a Switched Ethernet Network

3.2 IEC 61850-9-1 Traffic

IEC 61850 is a communication protocol for electrical utility substation automation where the effects of Ethernet frame latency must be taken into consideration. Consider the same switch as in the previous example, aggregating 61850-9-1 Sampled Value traffic to a single 100Mbps port. A configuration chosen to nearly saturate the 100Mbps link to the receiver is: a sampling rate of 128 samples per cycle (of 60Hz) and 12 individually connected “merging units”, each transmitting sampled value data frames. Each frame is 984 bits in length.

Assuming that sampling is synchronized across all twelve merging units, twelve frames will arrive at the switch every $1/(128 \times 60) \approx 130\mu\text{s}$. The best and worst-case latencies for a frame carrying sampled value data would be, respectively:

$$L(\text{best-case}) = (984 \text{ bits} / 100\text{Mbps}) + 5.2\mu\text{s} = 15\mu\text{s}$$

$$L(\text{worst-case}) = (984 \text{ bits} \times 12 / 100\text{Mbps}) + 5.2\mu\text{s} = 123\mu\text{s}$$

Note that the equipment receiving and processing Sampled Value data would need to be able to tolerate data delayed by as much as $123\mu\text{s}$ with a jitter of $108\mu\text{s}$ ($123\mu\text{s} - 15\mu\text{s}$).

3.3 Multiple switches

An Ethernet network installation will typically comprise multiple switches. Consider the case in which IEC 61850-9-1 Sample Value traffic from merging units on two edge switches is aggregated onto a single root switch, to which the receiver for all traffic is also connected.

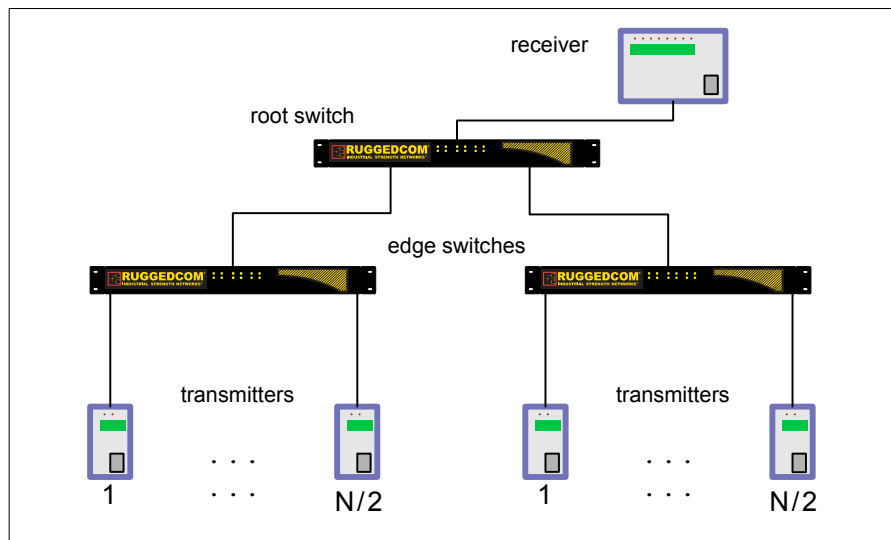


Figure 2: One root and two edge switches

Dividing the twelve merging units from the previous example evenly between the edge switches means that the queue on each switch grows to only $N/2$ frames. The queue on the root switch becomes the bottleneck, growing to N frames as, again, in the previous example. Since each edge switch begins forwarding as soon as the first frame has been received, the combined latency

Latency on a Switched Ethernet Network

increases only by the store and forward and switching times incurred by placing one additional switch in the path. Best and worst case latencies become, respectively:

$$L (\text{best-case}) = (984 \text{ bits} / 100\text{Mbps}) + 5.2\mu\text{s} + (984 \text{ bits} / 100\text{Mbps}) + 5.2\mu\text{s} = \mathbf{30\mu\text{s}}$$

$$L (\text{worst-case}) = (984 \text{ bits} \times 12 / 100\text{Mbps}) + 5.2\mu\text{s} + (984 \text{ bits} / 100\text{Mbps}) + 5.2\mu\text{s} = \mathbf{138\mu\text{s}}$$

Note that both latencies have increased by the time for one frame to cross one switch, whereas the jitter remains 108 μs , the same as in the previous example.

3.4 Mixed Network Traffic and Prioritization

Often it is not cost effective to dedicate a whole Ethernet infrastructure to a single application. In practice, many types of network traffic usually share an Ethernet. Critical network traffic is protected by marking it with a higher priority than other traffic. Using this technique to add general, low priority traffic to the network of the previous example, we find that in the worst-case, a high-priority Sampled Value frame could be delayed not only by other frames in its own queue, but also by one full-size frame in the low priority queue.

The best-case scenario remains the same; with no additional queuing effects, the latency remains 34 μs . In the worst case, however, it could be that a full-size frame has begun transmission from the low-priority queue just ahead of the reception of 12 frames in the high-priority queue. The worst-case latency becomes:

$$L_{MIXED} = [(12304 \text{ bits} + (984 \text{ bits} \times 12)) / 100\text{Mbps}] + 5.2\mu\text{s} + [(12304 \text{ bits} + 984 \text{ bits}) / 100\text{Mbps}] + 5.2\mu\text{s} = \mathbf{384\mu\text{s}}$$

which amounts to an increase in both the worst-case latency and the jitter by the store-and-forward latencies of two full-size frames.

If, furthermore, we were to add to this example a non-trivial wireline distance from one of the edge switches to the root switch - let us use the 100km mentioned earlier - the worst-case latency would increase by 500 μs to almost 900 μs . Consider the difference in magnitude of latency between this and the single frame case in the example of section 3.2, of 15 μs .

Latency on a Switched Ethernet Network

4 Types of Network Traffic

Other applications have different traffic patterns and delivery requirements, which can be characterized in the following ways:

- Real-time constraints
- Periodicity, message size and frequency
- Reliability requirements

Some examples of different types of IP traffic that might need to coexist on a LAN are:

- General IP traffic: A computer workstation is typically used for a variety of applications, which may generate and consume a wide variety of network traffic. For example, file and web transfer protocols like ftp and http can easily consume all available network bandwidth if they are not assigned a low priority relative to critical traffic.
- VoIP (Voice over IP): The commonly used G.711 audio encoder generates a 120-byte frame every 20ms. A good quality voice connection requires <50ms latency. Increased latency and some dropped frames may be acceptable while maintaining a usable connection.
- Streaming video: Video encoders often generate a data stream that has a consistent data rate over time, but alternates between large and small frames depending on the signal being encoded. Some loss (<5%) is acceptable for most codecs. As much as 5 seconds of latency might also be acceptable, depending on the application. A typical video stream might consist of an aggregate of 500kbps, roughly one-third of which might be full-size frames.
- Interactive video: Usually includes an audio stream like that used in VoIP. Latency requirements for interactivity are lower than for general streaming video (~150ms) and mean a greater susceptibility to lost frames, although some loss (<1%) is still acceptable.
- MMS (Manufacturing Message Specification): Messaging system for real-time data and supervisory control. Real-time updates are requested on the order of milliseconds. Frame sizes and frequencies are dependent on the type of MMS communication.
- IEC 61850 Sampled Value data, presented in the examples above, may be sampled at up to 256 times per cycle of 50 or 60Hz . Each Merging Unit transmits a 984-bit frame per sample period, and there will typically be several Merging Units on a network.

5 Conclusion

The basic factors contributing to message latency are the same on any switched Ethernet network, namely store and forward, switch fabric, wireline, and queuing latencies. The implementation and usage details of a network, including the number and layout of switches (both logical and geographical), the number and distribution of receivers and transmitters, their connection speeds and traffic patterns, all combine to determine how messages will be delayed when crossing the network. A simple model of the behaviour of a particular network, like the ones shown above, will help determine the parameters necessary to ensure reliable operation in practice.