Ethernet in electrical substations: Latency

Abstract

As a result of several international initiatives in the area of process automation in electrical substations, the issue of an international standard that would allow the interoperability of Intelligent Electronic Devices (IEDs) of different manufacturers was addressed. The IEC 61850 standard was Publisher. This standard defines the communication networks and systems that should be used for the automation of electrical substations. The standard details all the inter-IED communication related issues.

The chosen network technology for the interconnection of the different elements is Ethernet. The pros are obvious: it is a mature technology with a quite low cost and a performance (bandwidth and latencies) that can be suitable for its use in electrical substations as long as some questions are considered.

This paper discusses the latency of communications over Ethernet and proposes methods that can keep the latency predictable and controlled in the network of an electrical substation. This is necessary to meet the requirements imposed by IEC 61850 for the communication elements of an electrical substation.

Latency of a switch and IEC 61850

The IEC61850 [1] norm establishes the maximum value of the transfer time of a message according to its priority. For some of them, i.e. type I messages (trip and reconnect orders, amongst others), the maximum transmission time may not exceed 3 milliseconds.

The time of transference is defined as the sum of the processing time of the communications' stack of the message sender device (t_a), the transmission time of the network (t_b) and the processing time of the communications' stack of the message receiving device (t_c). Furthermore, IEC 61850 limits t_a and t_c below the 40% of the total transmission time [2]. This determines the latency of the network infrastructure of the substation, t_b, which may not exceed the 20% of the total transmission time. Thus the total time of communication must be less than 600 milliseconds (most restrictive case).

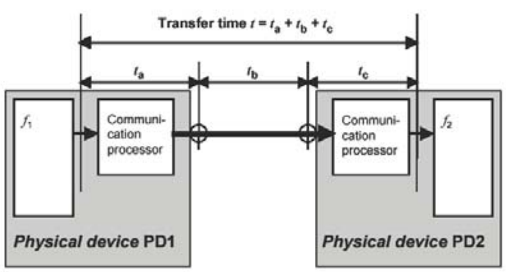

Figure 1. Transference Time of messages according to IEC61850 Standard

t_b includes the lags introduced by the communications' equipments (switches) and the propagation time of the frames through the communication medium (cables, fiber...).

The IEC 61850 functional certification for switches comprises transmission latency verification tests under presence of traffic of different priorities (e.g., high priority GOOSE messages in presence of low priority traffic) and some other tests to ensure there are no packet losses under different traffic circumstances and net-topologies.

The long way of a frame from the origin to the destination

There are two causes why an Ethernet frame needs a determined time to reach its destination when the frame is transmitted (we assume that the communication is taking place using a switch, so the physical medium is not shared, unlike when using a hub) and they contribute to t_b, the parameter mentioned above.

- **Physical propagation of the signals through the medium.** Whether an electrical medium (as a twisted copper pair cable) or an optical medium (plastic or glass optic fiber) is used, it takes a time proportional to the distance and inversely proportional to the propagation speed for the signals to propagate. E.g., a CAT-6 type cable has a propagation lag of 5 nanoseconds per meter. For the usual distances within a substation the total lag is less than a few hundred nanoseconds and it is negligible.

- **Time needed by the frame to "pass the through" the switch.** Present switches work on "store & forward" mode, that is, they receive the whole frame, they analyze it, they direct it to the corresponding port and they put it in the outgoing cable. The accumulated lag depends on the features of the switch, the speed of the involved ports, the amount of traffic at the switch at that moment… which may result in delays that may vary from a few microseconds to hundreds of milliseconds.

Propagation through the medium

In absence of any other network device, we can assume that the simplest network element would be a cable that would directly connect the origin and the destination. So the minimum time a frame would spend in the cable, t_b can be calculated. This time is equal to the total number of bytes of the frame plus prefixes (64 bits) multiplied by the time per bit (proportional to the transmission rate). The minimum inter-frame gap, which takes the time of 96 bits, must be added to that time. Thus a 64 bytes frame takes the following times depending on the transmission rate:

- 10Mbps: 672 bits x 100ns/bit = 67.2μs.
- 100Mbps: 672 bits x 10ns/bit = 6.72μs.
- 1Gbps: 672 bits x 1ns/bit= 672ns.

As analyzed in the preceding section, the propagation speed of the electric signals through a CAT-6 Ethernet cable is approximately 5 nanoseconds per meter. This is a very low value and it is not analyzed in this paper. It does have importance when considering trunk links with lengths of the order of kilometers, because in that case the latency of propagation will exceed the time the frame will spend in the cable.

Passing through the switch

The latency introduced by the switch must be added to the time calculated in the previous section. To that purpose it is necessary to previously explain the definition of latency that will be used from now on as detailed in the following paragraph.

For store & forward devices (such as the uSysCom switches), RFC1242 defines **latency** as the **time interval starting when the last bit of the input frame reaches the input port and ending when the first bit of the output frame is seen on the output port** [3]. This way the value of the latency is independent of the length of the frame. This is the standard definition of latency. If the time of the frame is added to the processing time of the switch, the t_b parameter of the previous section would be calculated. To do so the frame time (number of bits per frame x time per bit at the current bitrate) should have to be added to the latency calculated as it has been defined.

To get an idea of the complexity of a current Ethernet switch, the different stages that compound the switch should be analyzed in detail. These devices can be subdivided into three blocks:

- **Input block**. Each frame entering a port through a cable or fiber goes through the following stages:
    - PHY layer driver.
    - Media Access Control (MAC).
    - Input flow control management.
    - Input port queues.
    - Input port aggregation management.
    - VLAN filtering.
    - Get destination port.
    - Switching block entry queues.
- **Switching block**. All the input and output ports are connected with a shared memory switching matrix. The input frames are stored in memory with their destination port information and afterwards they are directed to these destination ports clearing the memory they occupied. The switching block also implements priority politics to put out the highest priority as fast as possible.
- **Output block**. This block is similar to the input block, in reverse. It can summarized as follows:
    - Frame processing, following the output rules.
    - Management and scheduler of output queues.
    - Transmission port aggregation port management.
    - Transmission flow control management.
    - Medium Access Control (MAC).
    - PHY layer driver.

It should not be forgotten that, usually, in a switch all this operations will take place sequentially in the specified order. This is why a perceptible increase of the latency of a frame when going through a switch should be expected depending on some factors as the following:

- Port traffic increase.
- Processing needs (MAC lists, VLAN access, priorities...)
- Port aggregation

Latency analysis for different implementations

In order to analyze the performance of its switches and obtain some conclusions about the most suitable configurations, uSysCom has carried out an analysis of the latency of its switches under various traffic conditions. For this purpose, some tests have been carried out using one or several frame flows that compete for the shared resources of the switch and assigning priorities per port and per VLAN tag.

The first one of these tests (dedicated ports) analyzes the behavior of the latency of the accessing traffic within the switch (traffic with origin and destination within the substation houses, that is, not using the trunk ports).

The other tests analyze the traffic going from the substation houses to the central building of the substation, that is, traffic going through the trunk ports. The selected topology in this case is a star topology, which minimizes the hops of the frames between their origin and the destination, thus its latency is smaller. The bandwidth of this links is not enough to send all the traffic of all the access ports with a 100% load, so it is an appropriate scenario to analyze the outcome of different priority politics in the performance of the switch.

These tests have been done with a 3SWT switch of uSysCom with 12 100BaseTx ports, 4 100BaseFx ports and 2 Gigabit Ethernet ports (SFP modules). A Smartbits analyzer of Spirent Communications was used as testing equipment.

Dedicated ports

Firstly a test was done, where several Ethernet frame flows enter through an access port of the switch and go out through another access port, so all the resources of the input and output ports are available for each flow (the flows are not mixed). The only resource that the entire frame flow uses is the internal switching matrix. But, since the internal architecture of the switch is non-blocking (wire-speed), the bandwidth of the switching matrix is big enough for to allow all the ports work at 100% load simultaneously.

The defined flows are the following:

- Flow from port 3 to port 4 of the switch.
- Flow from port 5 to port 6 of the switch.
- Flow from port 13 to port 14 of the switch.

Note that in this case, as each frame flow is independent of the others and as they do not share ports, no priority mechanism is being used since it would not bring any advantage. Using frame sizes of 64 and 1518 bytes and changing the volume of each flow between 50% and 100%, with flow durations of 10 seconds, the following came out:

Latency (µs)	P3 to P4 100BaseTx	P5 to P6 100BaseTx	P13 to P14 100baseFx
64bytes / 50%	3.0	3.5	3.1
64bytes / 60%	3.5	3.2	3.4
64bytes / 70%	3.0	3.0	3.1
64bytes / 80%	3.2	3.2	3.4
64bytes / 90%	3.2	3.6	2.8
64bytes / 100%	3.7	3.7	3.7
1518bytes / 50%	3.4	3.2	3.2
1518bytes / 60%	3.0	2.8	3.0
1518bytes / 70%	3.1	3.0	3.3
1518bytes / 80%	3.1	3.6	3.4
1518bytes / 90%	3.2	3.2	3.3
1518bytes / 100%	3.7	3.7	3.7

Table 1. Latency values obtained with separate flows.

As shown in table 1, the latency is below 4 microseconds for all the flows of the Ethernet ports and for all the analyzed traffic loads. During the test no frame was lost, so it can be assured that the access ports of the 3SWT switch are full wire speed.

Competing traffic-Priority per port

A quite usual communication network topology is represented in this case. There are 5 input flows coming into the switch through 5 different ports (access ports), and all of them 5 go to the output trunk port. Since all the ports are Fast Ethernet ports (100Mbps), with input traffic loads over 20% the output port will be saturated and some frames will be lost.

This configuration, with one trunk port and 5 100BaseTx access ports, may not be logical from the point of view of dimensioning of the trunk port, because, if the traffic load at all the ports could reach relatively constant and high levels, it would make sense dimensioning the bandwidth of the trunk port according to the "2+25%" rule (the bandwidth of two links plus 25% of the rest, in this case 2 x 100Mbps+0,25 x 3 x 100Mbps=275 Mbps)[4]. However the chosen configuration does make sense and would be useful in a low port-use scenario, this is why it was selected to illustrate the behavior of the switch's prioritization per port.

This way, using the priority per port mechanism, a particular traffic flow can be favored. A certain priority can be assigned to the traffic entering through each port and the switch will internally keep a percentage of its throughput for the highest priority traffic (if there is such traffic at that time).

So a high priority can be given to the traffic coming through a port (which communicates with a specially important device) and make sure this traffic will have place at the output port.

To check this, these traffic flows have been created (port 8 is the trunk port):

- Flow from port 3 to port 8 of the switch.
- Flow from port 4 to port 8 of the switch.
- Flow from port 5 to port 8 of the switch.
- Flow from port 6 to port 8 of the switch.
- Flow from port 7 to port 8 of the switch..

3SWT has three priority levels. All the access ports have the lowest priority level (1) except port 7, which has the highest priority (3).

The next figure shows the obtained results for a frame size of 64 bytes and a traffic load from 10% to 100%.

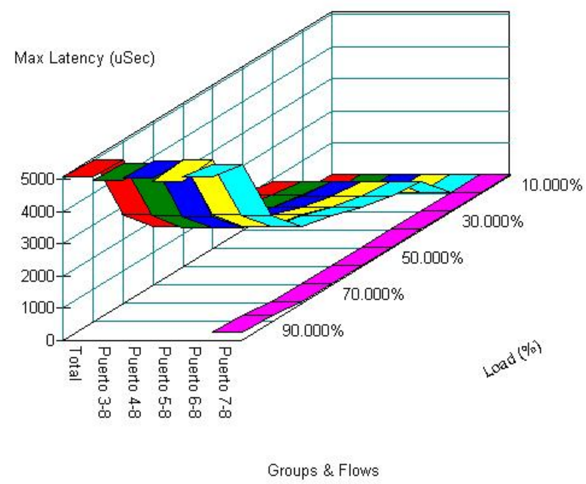

Figure 2. Latency for different priorities per port.

The latency for the traffic flow between port 7 and port 8 (high priority) remains stable around 16 microseconds, until the traffic-load reaches 80%. This is so because in this case the switch keeps the 80% of the throughput for high priority traffic, as it is shown in the figure of lost frames corresponding to the same case.

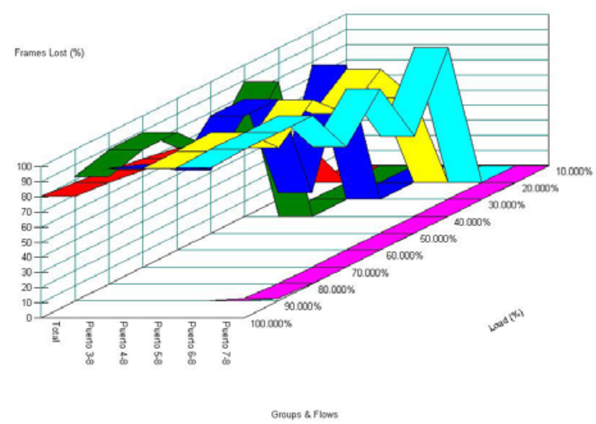

Figure 3. Lost frames for different port priorities.

The frame loss at the low priority ports starts when the traffic offered to the output port exceeds its capacity (5 x 20% = 100%). For traffic loads below 20% latency is similar to that of the high priority traffic. As the traffic load increases and some frames start to be discarded the value of latency is not meaningful.

On the other hand, the jitter at the high priority port remains very low as long as it is below the congestion zone (80%). On the contrary, since in low priority traffic-flows packets can pass depending on the resource availability at the time they arrive, the latency jitter is high from the point of maximum traffic-load without frame loss (20%) on.

Therefore, per port prioritization mechanism is very effective to prioritize traffic flows originated at a particular device, without distinguishing what kind of traffic it is. It should also be kept in mind that priority is only effective as long as high priority traffic is just a fraction of the available throughput [5].

Competing traffic-Priority per VLAN

Having traffic separated in different VLAN-s that can go to different control buildings is a usual case in the network of a substation. E.g., there can be a VLAN for the critical control information, another one for VoIP and another one for general purpose data.

Figure 4 shows an example of Ethernet in a substation. 3 VLAN-s have been defined. VLAN 1 represents a control VLAN that goes from the command post of the central building to the switching relays of the substation houses. VLAN 3 connects the IP phones from the different houses that are used to communicate with the central post and VLAN 2 groups the existing general purpose traffic of the network.

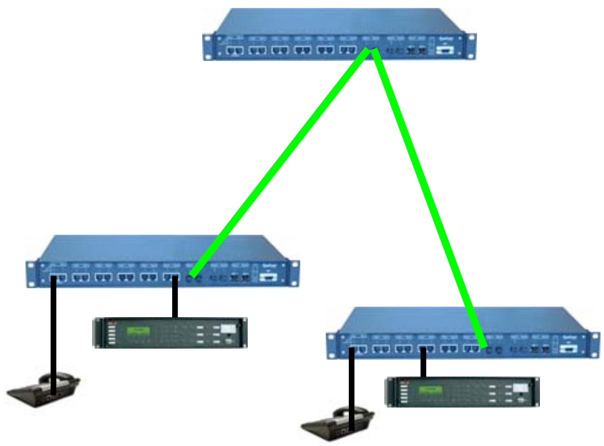

Figure 4.Example of a substation network.

For testing this scenario several flows, included in three VLAN-s, have been defined: One of them has low priority (VLAN 1), another one has medium priority (VLAN 3) and the last one has high priority (VLAN 2). To be precise, these are the VLAN-s (port 6 is the trunk port):

- From port 1 to port 6, with VLAN 1.
- From port 2 to port 6, with VLAN 2.
- From port 3 to port 6, with VLAN 2.
- From port 4 to port 6, with VLAN 3
- From port 5 to port 6, with VLAN 3.

The following output bandwidth is kept for each priority level:

- High priority: up to 60%.
- Medium priority: up to 30%.
- Low priority: up to 10%.

Port 6 (painted green at figure 4) is the trunk port that connects the substation houses with the central building of the substation, and the other ports are access ports. The next figure shows the results for a frame size of 64 bytes and a traffic-load increasing from 10% to 100%.

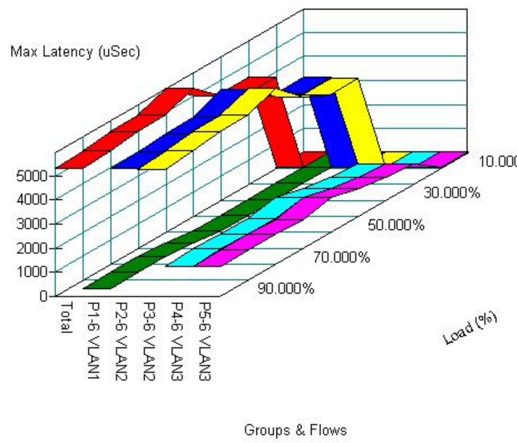

Figure 5. Latency for different priorities per VLAN.

As shown in figure 5, the latency of the frames of VLAN 1 is low and remains around 16 microseconds until the traffic load exceeds 60% of the output port's bandwidth, which is the maximum guaranteed bandwidth for the high priority traffic. It is remarkable that the latency of the traffic flows of VLAN 3 (medium priority) is sensibly smaller than that of the traffic of low priority (VLAN 2), even if, from 20% of traffic on, a significant number of frames is lost, as shown in figure 6.

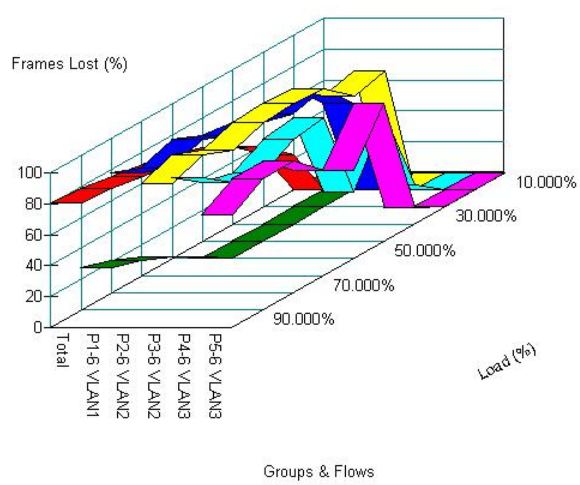

Figure 6. Lost frames for different VLAN priorities.

It has also been checked that, for high priority traffic, before the links are saturated, the latency jitter is very small (of the order of 10% of the total latency).

It can also be noticed that, with this configuration, there is almost no bandwidth available for low priority traffic as soon as the higher priority traffic grows. Thus the conclusions of the preceding case also apply here.

Some usual troubles

Clock offset

If a switch sends frames with a higher clock speed than the clock speed of the destination device, the destination device will either have to discard or store the frames. The longer the sent high speed frame sequence, the bigger the needed storage space for the excess frames. This is a trouble when using Gigabit Ethernet interfaces, since a lot of information is transmitted and the required storage space can be very important [6].

The norm 802.3z specifies that the clock used for Ethernet interfaces may have a tolerance of ±100ppm (parts per million). This means that, in at worst, a switch would have to receive the packets at one port with clock speed +100ppm and send them through another port with clock speed –100ppm speed.

As an example, if we consider 64 bytes long frames, with a prefix of 64 bits and an interframe gap equivalent to 12 bytes, these are the results for a Gigabit interface:

TX frames (exact clock): 1,488,095.238 fps.

TX frames (clock +100ppm): 1,488,244.048 fps.

TX frames (clock –100ppm): 1,487,946.428 fps.

This means that the switch would have to store 297.619 frames per second (approximately 19 Kbytes/second) for each pair of Gigabit Ethernet interfaces working at 100% capacity. This can be a problem when long bursts occur, since output buffer overflows could take place in a few seconds.

This problem an be eliminated by dimensioning the system in a way that Gigabit Ethernet interfaces will not be under traffic congestion extendedly, or by duplicating capacity using mechanisms such as port aggregation.

Traffic prioritization criteria

As network resources are finite, it is necessary to administrate them carefully so the applications that really need priority are the ones that get it. A network management maxim states that “if everyone uses high priority, no one actually gets it” [7]. Thus it is necessary to classify the traffic according to its characteristics and assign the corresponding priority. Obviously, this scheme has to be studied for each specific case.

The 3SWT equipment of uSysCom offers three different priority levels. For an electrical substation environment, this could be a possible priority implementation:

- Maximum priority: Network and equipment management traffic (SNMP...).
- Medium priority: High priority application traffic (functions requiring minimum latencies and functions for which packet loss is not acceptable).
- Minimum priority: The rest of the traffic (it will be sent as long as resources are available).

The need to assign the highest priority to the network management traffic is easy to understand. If any problem happens, the network administrator should be able to access and reconfigure the equipments. It should also be kept in mind to adjust the bandwidth that is reserved for each priority level, so the bandwidth is not to be wasted. In most cases, management traffic has a low volume.

The intermediate priority level can be assigned for the traffic of applications that are especially important (e.g., in a substation, type I messages). If there was priority traffic with a high data volume the available bandwidth should be increased (upgrade to Gigabit Ethernet, port aggregation), since, if priority system is to give the desired results, it is necessary that priority traffic is not the majority of total traffic.

Finally, the rest of the traffic is the one that is generated by non-latency-critical applications that have higher layer protocols for lost frame recovery.

References

[1] “IEC61850-5, Communication networks and systems in substations. Part 5, Communication requirements for functions and device models”. IEC, 2003.

[2] “IEC61850-10, Communication networks and systems in substations. Part 10, Conformance Testing”. IEC, 2005.

[3] “RFC1242, Benchmarking Terminology for Network Interconnection Devices”, IETF, 1991.

[4] Dooley, K. “Designing Large-Scale LANs”, O’Reilly, 2001.

[5] Ma, A. “Multilayer Switch Testing and Analysis”, Spirent Communications, 2003.

[6] “Smartbits System Reference”, Spirent Communications, 2004.

[7] Seifert, R. “The Switch Book”. John Wiley & Sons, 2000.

Aitor Arzuaga Munsuri

Born in 1975, Aitor is a Telecommunication Engineer in the specialty of Radio Communications graduated from the ETSIIT of Bilbao.

During the last 7 years he has participated and leaded the design of communication products using diverse technologies, both wireless (GSM, GPRS, GPS, Bluetooth, Wireless LAN) and wired (Ethernet, field buses). He is specialist in communication electronics design.

Currently he is Technical Manager of the networking products area of uSysCom. He has published several articles on the application of communication technologies for industrial environments and robust communication equipments.

uSyscom designs and manufactures cutting edge communications equipment for utility and industrial applications. uSyscom is part of the ZIV Group and its headquarters are located in Zamudio, Bizkaia (Spain).

See www.usyscom.com and www.ziv.es for more information

Since 1999, ZIV USA Inc. has been providing local technical support for ZIV customers. ZIV USA Inc. is located in the Chicago area with several representatives around the USA and Canada.

See www.zivusa.com for more information.