Ethernet with Time Sensitive Networking Tools for Industrial Networks

Csaba Simon, Miklós Máté, Markosz Maliosz and Norbert Bella

***Abstract* — In currently deployed networks the time critical and/or real time traffic is sent over dedicated networks, requiring the operation of a separate infrastructure. This is especially true for Industrial Networks, which use technologies and protocols that are designed particularly for that purpose. The IEEE 802.1Q Time-Sensitive Networking (TSN) task group introduced a set of standards by defining QoS mechanisms, also known as TSN features, so that standard Ethernet networks could provide precise timing for critical flows. We have implemented two mature TSN features, frame preemption and time gated queuing, in a simulator, and on multiple network topologies we have evaluated the end-to-end delay and packet delay variation as the main QoS metrics and important design considerations in industrial networking setups. Our simulation results have shown that the QoS guarantees provided by TSN are strong enough for industrial use cases, but we have also identified some design and configuration pitfalls that TSN-adopters need to be cautious about.**

Index Terms* — *Industrial Networks, IoT, Ethernet, Time-Sensitive Networking

I. Introduction

The latest developments in manufacturing technology and the globalization of world economy increased the competition among manufacturing sites at global level. The spread of smarter devices (machines) and the decreasing cost of computational resources enabled the innovation of manufacturing processes. Industrial actors have to follow the rapid changes in customer demands and this high flexibility must be supported by the plant infrastructure. The Industry 4.0 initiative, catalyzed by the German Federal Ministry for Economic Affairs and Energy [1] summed up the digitization trends in manufacturing. Industry 4.0 has also been acknowledged by the European Union (EU) to set a framework for the required research and development actions to be taken along the road of the digitization of manufacturing [2]. Since then the Industry 4.0 has been enumerated among the foundations of the European Commission's industry-related initiative of the Digital Single Market, which forms the industrialization policies in one of the world's leading industrial ecosystem [3].

The realization of the Industry 4.0 goals towards the digitization of the industry also puts a great pressure on network solution providers. Traditionally the networks connecting the process controllers, machines and sensors have been using specialized equipment and were expensive and lacked adaptability. Whereas industrial networking and Industrial Ethernet solutions in particular are available for many particular use cases, the interaction and cooperation among these solutions is not solved [4]. Moreover, the adaptation of these existing solutions to some of the needs is cumbersome, too [5]. As digitization of the real time manufacturing process has strict requirements in terms of timing and data rate, the provision of proper transport/networking solutions is crucial to its success. There is a growing consensus among actors in the field of industrial networking that standard Ethernet should be used to provide the much needed one-size-fits-all solution.

Using Ethernet in an industrial environment is nowadays realized by *Industrial Ethernet* that is the use of Ethernet with protocols and/or modified MAC (Media Access Control) layer to provide determinism and real-time control (see Section II). Our paper promotes an emerging alternative based on standard Ethernet with the enhancements added by Time-Sensitive Networking features. The use of standard Ethernet (or "vintage Ethernet", as referred to in [4]) for industrial networking offers the promise of cheap operation, supported by the economies of scale resulted from the huge installment base and mass-scale production, large pool of specialists and low training costs, a well understood behavior tested in various deployment scenarios and over all phases of its lifecycle. As matter of fact, nowadays Ethernet is the de-facto single standard for layer 2 technology in enterprise networking and data centers [6][7], and a powerful alternative in the aggregation and access domains of the telco networks [8]. Ethernet also enables convergent networking, where various types of traffic are transported over the same infrastructure. In the context of industrial networking this means that low priority (e.g., best effort) traffic is carried on the same transport where high priority control traffic is forwarded. On top of that, due to this property, the use of standard Ethernet instead of various Industrial Ethernet standards also makes these deployments future proof, because it enables any new applications and services operating on this common infrastructure.

All the above considerations highly motivate both business decision makers and network architects to adopt Ethernet for industrial networks. Still, certain real-time manufacturing applications are safety-critical and have strict Quality of

Manuscript submitted May 12, 2017, revised June 20, 2017.

The authors are with the Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, Budapest, Hungary, 1117 Budapest, Magyar Tudosok krt. 2. (e-mail: {simon|mate|maliosz|bella}@tmit.bme.hu)

Service (QoS) requirements in terms of availability, packet loss and delay, not supported by the standard Ethernet. Current installed base and current product lines of the vendors offer further arguments against this change. Nevertheless, latest evolutions in Ethernet standard expected to be embraced by both vendors and system integrators in the near future strengthens our expectation that the networking and automation industry are at a turning point, business-side demands and technical expectations making this change to happen. Looking at the standardization activities concerning Ethernet, the most relevant advances are made within the IEEE 802.1 Time-Sensitive Networking (TSN) task group [9]. TSN mechanisms address all critical aspects of industrial applications requiring strict QoS, synchronization, reliability and deterministic delay (see Section III for further details). Note that several use cases are expected to deploy TSN mechanisms (5G, vehicular, IoT, etc.), but in this paper we keep our focus solely on industrial networking. Since TSN features are defined as amendments to the main IEEE 802.1Q standard [10], major vendors will support it in the near future (e.g., in industrial networking nodes by TTTech [11] or Cisco [12]) and industrial networking solution providers (e.g., National Instruments [13], Belden [14]) will propose integrated solutions to manufacturers.

Based on the above argumentation we predict that in the near future industrial networking will be based on standard Ethernet (which also includes TSN features). Knowing that industrial traffic has complex requirements the introduction of standard Ethernet in industrial networks is not straightforward, involving more engineering effort than just replacing current networking nodes. Before such change happens we should understand how the Ethernet network will accommodate the industrial traffic.

We use a network topology and traffic mix specific to industrial networks to illustrate the way how standard Ethernet can support industrial applications. Building on the results of our simulation experiments we explain some issues that might occur in Ethernet based industrial networks, which the designers of such networks should be aware of. One goal of our paper is to show by simulations the efficiency of using different TSN features to protect the express traffic against the low priority (best effort) traffic. Another goal of our paper is to highlight the interaction between multiple express flows, because TSN features will protect the express flows from the low priority traffic, however using them it can result in race conditions between the frames from different express flows causing delay variations.

The rest of this paper is organized as follows. Section II gives an overview of the most common Industrial Ethernet solutions. Section III summarizes the two TSN features (preemption and time-gating) we used in our simulations. Section IV presents our simulation setup and the results of the simulations, and in section V we draw the conclusions.

II. INDUSTRIAL ETHERNET SOLUTIONS

In the field of industrial control and monitoring technology there has been an ongoing transition from vendor-specific and other specialized systems to Ethernet. Ethernet have evolved way beyond its original CSMA/CD-based bus architecture in the last couple of decades. Its new features and increased speed make today's switched Ethernet a reliable, cheap and versatile telecommunication medium. With an array of proprietary extensions it is possible to deploy it in a factory.

There are three levels of quality requirements in industrial networks. The first level is when there is only best-effort traffic: the only requirement is a low packet loss rate, but there are no timing requirements. The second level is the soft real-time systems: here the requirement is that the packets must be delivered within 10 ms with the lowest delay variance possible, and no out-of-order delivery is allowed. The third level is the hard real-time systems: the packets must be delivered within 1 ms, and there is even less tolerance for delay variance than on the previous level.

Industrial Ethernet is the common name for network technologies that are Ethernet-based solutions for industrial control, monitoring and automation systems. They include extensions to standard Ethernet that enable deterministic delay, high reliability and real-time response capability. These equipment are also hardened for the potentially harsh environmental conditions of a factory, but the hardware specifications are not in the scope of this paper. In the followings we provide a non-exhaustive overview of Industrial Ethernet technologies.

Perhaps the most widely known Industrial Ethernet solution is PROFINET (PROcess FIeld NETwork) [15]. The PROFINET IO system consists of one Controller node, and a number of Device nodes connected by Ethernet links. PROFINET IO RT provides soft real-time communication with cycle times between 5 ms and 10 ms. This is essentially best effort with optimized software stack for fast processing. Without time synchronization the frames sent by the slave nodes may collide; thus, star topology is recommended. PROFINET IO IRT is an isochronous, hard real-time variant. With precise time synchronization short cycle times can be reached; the specification supports 1 ms and 250 us. When the IRT data does not fill the whole cycle, the remaining time can be allocated to RT or other TCP/IP data. Both RT and IRT supports line, star and ring topologies, but the timings must be tuned according to the number of devices and cable lengths.

EtherNet/IP adapts the Common Industrial Protocol (CIP) to Ethernet [16]. Its communication model is based on a producer/consumer model, a message is only delivered to the nodes that subscribed to it. This is, however, implemented as broadcast transmission and filtering at the receivers. The nodes are aligned in a ring topology; the ring manager node verifies the connectivity every 400 microseconds by polling the nodes in the ring, and reconfigures the network if it detects an error. The protocol uses TCP connections between the nodes for configuration and management functionalities, while the I/O traffic is sent over UDP.

CC-Link IE was originally developed by Mitsubishi Electric [17], but later it was released as an open standard maintained by the CC-Link Partner Association. It has application profiles for communication between process controllers, field I/O, and motion control. The nodes form a ring topology of at most 120 nodes, and their communication pattern is governed by token passing supervised by a master node. The physical layer used is 1 Gbps Ethernet on fiber or copper, but the protocol uses a special MAC. Some versions of CC-Link also support star and bus topologies in addition to the ring.

Sercos III is the third generation of the Sercos automation bus [18]. It is based on standard Fast Ethernet (100 Mbps) with EtherType 0x88CD, and the network topology can be either a bus or a loop. The communication is cyclic, the length of one cycle can be set between 31.25 us and 65 ms, depending on the number of devices. Part of the cycle is reserved for the real-time traffic of Sercos, while the rest is freely available to other traffic, such as TCP/IP. Sercos also synchronizes the devices in a network to the master clock with 1 us accuracy.

As a summary we can conclude that these protocols use a master-slave hierarchy and periodic polling to achieve deterministic delays and fair channel arbitration. The specifications also pin the speed of the underlying Ethernet interface, which is good for determinism, but it can limit the scalability of the network.

Based on the publicly available market research data over the last decade among all the Industrial Ethernet technologies the PROFINET is among the most successful solutions and has the most dynamic growth [19][20]. Due to the popularity of this protocol we will use it as a reference networking scenario in this paper, as detailed later in Section IV.

III. Time-Sensitive Networking

Ethernet is envisioned to be used as single standard in Industrial Networking, however the current standard Ethernet without the TSN features does not offer QoS guarantees required by the industrial applications. Historically the QoS support in Ethernet has been mostly limited to priority queueing. The VLAN-tag defined in IEEE 802.1Q includes a 3-bit field called Priority Code Point (PCP), and IEEE 802.1p defines the meanings of these 8 priority levels. The transmission selection logic in an Ethernet switch uses separate queues for each priority, and a priority-based selection algorithm [10]. A dedicated task group (TG) called Audio Video Bridging (AVB) was established to expand the IEEE 802.1 bridging features for better QoS support, which later was renamed to become the IEEE 802.1 Time-Sensitive Networking (TSN) Task Group to extend the scope to industrial and automotive fields [9].

The TSN TG is specifically working on enhancing the standard Ethernet bridging to support time-critical traffic by defining standardized solutions for QoS mechanisms to reach precise timing for the priority traffic in Ethernet networks [9]. The task group is still active and more mechanisms are expected to be defined in the following years. In the remaining part of this section we succinctly present the mature TSN features. Later in our work we used only those that we considered that support the industrial applications.

Time synchronization between the switches is an important feature of TSN. The Precision Time Protocol (PTP) as defined in IEEE 1588 was adapted to Ethernet in IEEE 802.1AS, and TSN extended it in IEEE 802.1AS-Rev [21]. This protocol can achieve sub-microsecond accuracy, which is sufficient for most TSN applications. Synchronization might not be perfect and applied a stochastic model to examine the effects of this imprecision.

IEEE 802.1Qcc Stream Reservation defines the layer 2 level reservation procedures [22]. In typical industrial networking scenarios the resource reservation for communication is done before the start of the industrial applications. Therefore we are not detailing this standard in this paper.

IEEE 802.1CB proactively replicates packets to offer redundant transmission and it manages the handling of redundant packets to make this process transparent to the application [23]. This Frame Replication and Elimination for Reliability (FRER) mechanism increases the packet delivery ratio, which is useful in certain control protocol scenarios, if the recovery from an equipment failure is unacceptable. FRER essentially provides 1+1 protection for streams. In most of the industrial networking scenarios this requires the deployment of a second “shadow” network, which might be feasible for specific use cases. In our investigations we focus on such optimized deployments where only single streams are sent in the network.

A new mechanism of TSN that is very different in nature compared to other ones standardized by the IEEE 802.1 work group is the frame preemption, which is standardized in IEEE 802.1Qbu [24] and IEEE 802.3br [25]. With this mechanism the transmission of a lower priority frame can be interrupted in favor of transmitting a higher priority one. The transmission of a frame on a point-to-point Ethernet link was traditionally thought of as an atomic operation. Interrupting it was seen as violation of a basic, yet unwritten Ethernet principle by many professionals. Nevertheless, the power of preemption for providing bounded delay variance is very convincing.

Preemption works as follows. The 8 priority queues are divided into two groups: the highest priority queues are labeled as express, and the rest are preemptable. When an Ethernet frame arrives in an express queue, and a preemptable frame is being transmitted, the transmission is interrupted, the remainder of the frame is put aside, and the transmission of the express frame is started. When the transmission of the express frame is finished, and no other express frame is waiting in the queue, the transmission of the preempted frame is resumed. One transmission can be interrupted several times.

By using preemption the delay and delay variance of time-critical traffic can be greatly reduced without waiting for the end of the transmission of a potentially large frame ahead of an express frame. There are limitations to the efficiency of preemption though. The transmission can only be interrupted on byte-boundaries. The MAC also must send a CRC inserted into the tail of the frame, and wait an inter-packet gap (IPG) before a new transmission can be started. This means that there is always a non-zero preemption latency. The minimum size of an Ethernet frame is 64 bytes; thus, the express frame suffers higher preemption latency if it arrives at the beginning or at the end of the transmission of the preemptable frame.

The worst case happens when the preemptable frame is too small, and cannot be cut into two pieces of at least 64 bytes each [24]. Note that express frames cannot be preempted; thus, the interference between time-critical flows cannot be handled with this technique.

The other major TSN mechanism for express traffic is the time-aware shaper standardized in IEEE 802.1Qbv [26]. It introduces a timed gate for each of the 8 priority queues, the opening and closing of these gates is governed by a predefined gate control schedule. When a gate is closed, the corresponding priority queue is barred from the transmission selection algorithm.

If the arrival times of the express frames at a switch are known, then this tool can completely eliminate the latency caused by interfering traffic. In anticipation of the express frame the gates of all other priority queues are closed, and when the express frame is expected to have left the switch, the gates can be opened again. This sounds simple at first, but there are several pitfalls when setting up the gate timings. If the express frame arrives outside of its expected interval, it is not protected from the interfering traffic. If the express frame doesn't arrive at all, because the flow hasn't started yet or it already ended, the other flows were stopped uselessly.

Most of these problems can be traced back to incorrect time synchronization among the devices in the network. Time-gating can only function properly, if all the devices are synchronized to a master clock. Time gating can be protected from the synchronization uncertainty by increasing the expected intervals with the upper bounds of the time synchronization errors.

The above two TSN mechanisms can be combined into a so-called protected window with guard band. In this mode the preemption is triggered by the gate timing instead of the arrival of the express frame. In anticipation of the express traffic its gate is opened, and at the same time a preemption signal is sent to the MAC. When the express frame arrives, the way is already cleared for it. The advantage of this method over simple time-gating is higher link utilization: without preemption support the transmission of a lower priority frame cannot start, if it won't finish before its gate closes.

IV. SIMULATIONS

A. Simulation Scenarios

We prepared several simulation experiments to find answers to the questions asked in Section 1. The network topology of our simulations is based on the topology observed in a real PROFINET IO RT deployment investigated by Ferrari et al. [27]. The core switches are connected in ring topology, and the IO devices are connected to the core switches via cascaded buses by lines of drop switches. Maximum 9 IO devices are cascaded over one bus, but multiple such buses may be plugged into a core switch. All devices are controlled by one IO controller, which is directly linked to one of the core switches. For performance reasons we only simulated part of this topology, but taking care to keep all the important node and traffic types. We will refer to this topology as the cascaded one. Starting from this topology we also created another two topologies to better understand the performance of TSN.

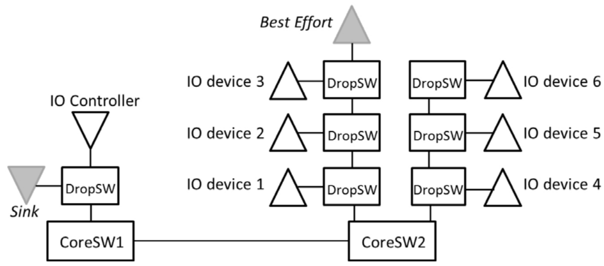

Fig. 1. The simulated network with cascaded topology

Our second topology is a slightly modified version of the cascaded one. In this topology the IO devices are all connected to their respective core switches directly, without forming a bus. This topology corresponds to the usual switched Ethernet topologies and not to the usual industrial topologies. We will refer to this one as the star topology.

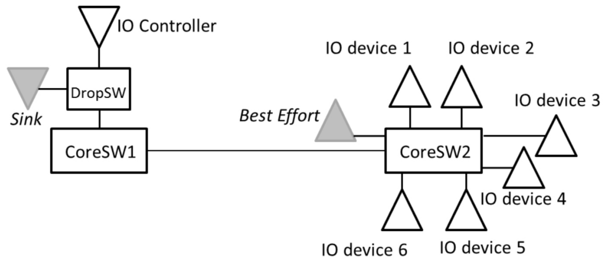

Fig. 2. The simulated network with star topology

In our third topology a series of IO devices are connected into a ring, one device per each switch, forming a ring topology. In practice multiple devices might be connected to the switch. The idea behind this arrangement is that there are few devices attached to the core switches and the ring is expanded instead.

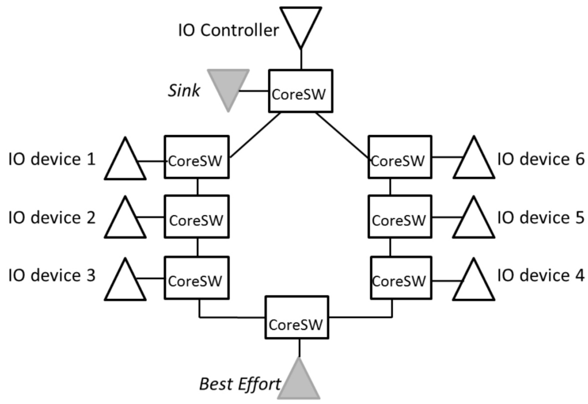

Fig. 3. The simulated network with ring topology

This topology is formed with a different mindset compared to the star topology, because the latter tries to channel as many devices possible to the core switch. The redundancy that the ring provides is very important in factories, where high reliability is absolutely required. Note that in the cascaded topology if the bus is detached from the core switch, all the devices from it are disconnected from the controller.

Fig. 1, Fig. 2 and Fig. 3 present the logical structure of the three topologies implemented in our simulation tool. In a large real deployment there would be hundreds of devices in a network, but we had to limit our simulations to 6 IO devices and one low priority traffic generator (the Best Effort source that send frames to the Sink). According to TSN terminology the end devices are called *end stations*, but in our industrial setup they are denoted as *IO devices*. In all of the three examined topologies we could examine the conflict situations when multiple control flows race for the same output link, causing QoS degradation (packet delay variation).

In all three networks the core switches are connected by 100 m long 1 Gbps links, and all the other links are 10m long and have 100 Mbps link rate. The serialization delay of 1bit over the 1 Gbps link is 1 ns and over the 100 Mbps link is 10 ns. The propagation delay is 5 ns/m.

In the case of cascaded topology all the IO devices are linked to a small switch, called drop switch, which in practice is integrated with them. In both cascaded and star topologies the sources send the frames at around the same time with only a small time difference (in the order of ns). Thus, the flows originated from devices that are at equal distances from the core switch will arrive at around the same time and with precise timing this small time difference makes the servicing order deterministic between the frames of these flows. However, with imprecise timing these frames are in race condition for the output link.

When designing the traffic model we attempted to recreate the worst-case situation. The control traffic is Constant Bit Rate (CBR) traffic, with 1 ms cycle time (the time between two successive frames). This 1 ms value provides hard real time assurances in Industrial Ethernet [28]. Unless stated otherwise, we used frames with MTU (Maximum Transmission Unit) of 1500 bytes, because that is the maximum frame size allowed by Ethernet. These large frame sizes are the worst case, because smaller frames are less prone to collisions and have lower serialization delays over the same link rates. Best Effort traffic has 1500 byte frames, but are sent with a different cycle time (1.2 us) to have race conditions with different control flows during the simulation time, thus we could observe every possible interference pattern between the high and low priority traffic. The control traffic between the industrial devices and the controller is the prioritized one, which is called in TSN terms the express traffic. The other, low priority traffic is considered to be best effort (BE) traffic. The switch model in the simulator is a generic switch with infinite backplane switching capacity and with switching delay value of 1.5 us, which models a fast enterprise switch [29].

We required frame-level analysis for the experiments; therefore the simulation environment must be a frame-level tool, providing the Ethernet stack, preferably extensible in a modular fashion. The OMNeT++ tool [30] with its INET framework [31] satisfies this requirement. The existing Ethernet model is easily extensible. We implemented the TSN features for our investigations, such as time gates and preemption support.

B. Simulation Results

Our goal is to evaluate the effectiveness of the usage of selected TSN features by analysing the end-to-end delays of the frames in the network. In order to assess the delay variation we monitored the Packet Delay Variation (PDV). PDV is the commonly used term for packetized traffic, even if in our case at layer 2 we observe the delay variation of frames. The PDV is computed as the difference between the highest and lowest end-to-end delay value:

$$PDV = \max(\textit{end-to-end latency}) - \min(\textit{end-to-end latency}) \quad (1)$$

Also design guidelines can be inferred from the observed results, e.g., regarding the timing configuration of the time gated queues.

1) Assuming Precise Traffic Source Timing

For the first set of experiments the traffic sources send periodic data in cycles, and here we assume that the sending is precise, thus exact timing without any variations is applied in the simulator.

In our first experiment the sending time of the express frames are set such as they will leave the ring switch in a deterministic way, resulting in fixed end-to-end delays (i.e., zero PDV).

Fig. 4 plots the end-to-end delays of the frames of all the 6 express flows with different colors. Each point of the chart corresponds to one express frame, and for the each flow the end-to-end delay is the same for all of its frames. The chart shows a repetitive pattern, because all the frames during each cycle are scheduled to be sent at around the same time, and the small time differences yield deterministic arrival to CoreSW2.

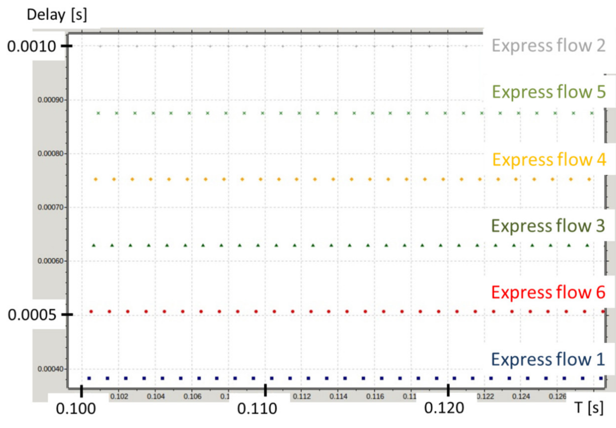

Fig. 4. Six express flows in a network with star topology

When BE traffic is enabled and there are no TSN mechanisms deployed the frames within express flows suffer different end-to-end delays, resulting in non-zero PDV (see Fig. 5). The reason for this is that while the BE frames are served by the switch, the express frames need to wait in the queue, and they are therefore delayed. When an express frame is delayed, another express flow can take the “place” of this delayed frame. On the figure this looks like the frame in question is “shifted”, as the three arrows show in zoomed in part. It can also be seen that BE has a repetitive nature, but with a different cycle time and because of this it delays express frames at different positions in the back-to-back train of the 6 express frames, resulting in different end-to-end delays.

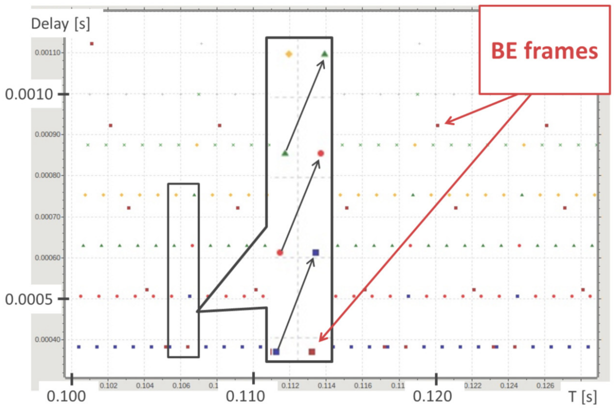

Fig. 5. The effect of BE frames on express traffic (star topology)

The effect of BE flows heavily depends on the topology and original timing of the express flows. We illustrate this in Fig. 6, where the pattern is slightly different compared to Fig. 5, because the arrival of the two sources closest to the CoreSW2 get through the switch before the BE frame would affect them. Thus, only the remaining four express frames have non-zero PDV.

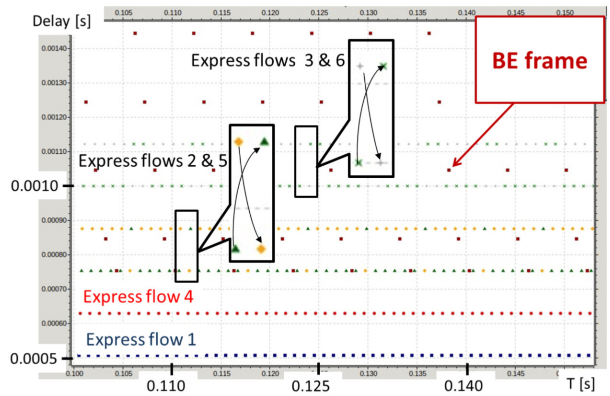

Fig. 6. The effect of BE frames on express traffic (cascaded topology)

In order to avoid the delays caused by the interfering BE frames, first we applied the preemption mechanism (see Fig. 7). The numerical results are shown in Table I. The PDV = 0.15 us for star topology, which corresponds to the worst case (i.e., largest) preemption delay. Note that for the ring and the cascaded topologies the PDV is 123.04 us, a value that is three orders of magnitude larger than the preemption delay. The reason for this is that in the worst case the preemption delay is large enough to "shift" the preempted frame behind the express frame from a different control flow, and the PDV is

$$PDV = \textit{preemption delay} + \textit{serialization delay(the other express frame)} \quad (2)$$

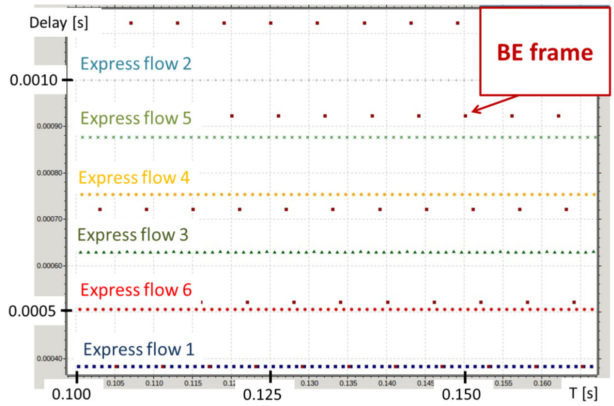

Fig. 7. The BE and express flows in a star topology network with activated preemption feature

Note that in this case the express frame size is the largest one, equaling 1500 bytes, thus, we measure the highest PDV possible. These experiments show that the choreography of the express frame arrivals is fragile: even very small delays may result in large PDVs. During the design phase it is not enough to evaluate the effect of race conditions at a single hop. Uncertainties induced by low priority traffic at a given hop might result in more race conditions farther away along the path. This can be avoided if we increase the gap between the arrival of the express frames.

TABLE I
HIGHEST PDV VALUES FOR ALL THE SCENARIOS WITH EXPRESS FRAMES OF 1500 BYTES

Topology	Express only (us)	Both express and BE (us)	Preemption (us)	Time gating (us)
Cascaded	0,00	123,04	123,04	0,00
Star	0,00	123,04	0,15	0,00
Ring	0,00	212,47	123,04	0,00

We also applied the time gated queuing to all three scenarios (see the last column of Table I). We calculated and applied the proper time gating configuration for each switch to let the frames of express flows through the network without getting into race conditions, and the resulting PDV is 0 for all topologies. Note that this perfect result was achieved in a setup without any variation of the source sending intervals, or of the switching delay values. This shows that in a well synchronized network time gating is able to protect the express traffic from the potential uncertainties caused by the BE frames.

a) The effect of link rate of the IO controller

In the followings we analyse the effect of the link rate at the controller (e.g., the output link of CoreSW1 towards the IO controller in Fig. 3). Let us start with the network arrangement investigated above. Fig. 8 shows the frame positions within a cycle as seen at the output links of the input of the CoreSW2, at the output of the CoreSW2 and at the output of CoreSW1 in the case of ring topology. The frame sizes are 1500 bytes and the source sending times are set in such a way that they arrive back-to-back to the CoreSW2. The input link rates of CoreSW2 and the output link rate of CoreSW1 are 100 Mbps, the core link between the core switches is 1 Gbps.

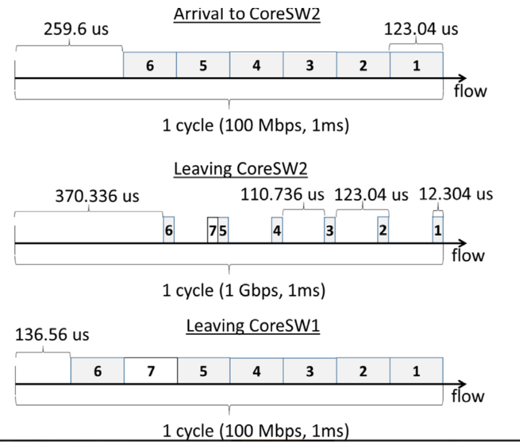

Fig. 8. Frames of racing express flows seen at different hops of the star topology

Even if the frames arrive back-to-back at CoreSW2, since the output link rate of CoreSW2 is 10 times faster and subsequently the serialization delay of each frame is 10 times shorter, there will be gaps between them. In this gap 9 similar express frames, which come from a different direction, may get inserted. For the sake of easier understanding we illustrated the insertion of only a single additional frame 7 right after frame 5 at the output of the CoreSW2. On the 100 Mbps controller link the serialization delays of each frame increase again to the 123.04 us value, and frame 7 will shift frame 6. This results in a PDV of frame #6 equal to this serialization delay.

Now let us evaluate the effect of increasing the link rate of the controller link to the link rate of the core link. Then both links rates will be 1 Gbps, the serialization delays of the frames will not change at the core switch, thus the gaps between the frames will remain the same. The possibility for other frames to be inserted will still be open (e.g., the frame #7 can be inserted between frames #5 and #6), but they will not cause any further delays. This means, that the PDVs of the observed frames will not be further increased.

We made a simulation experiment where the control link rate was 1 Gbps to show that the pacing of express frames remain unaltered by the insertion of a new express frame, as shown in Fig. 9. For this experiment we used the cascaded topology, in which we increased the control link rate to 1 Gbps and we kept only two express flows. We scheduled their arrival at CoreSW2 back-to-back, as it can be seen in the left hand side of Fig. 9. Then we sent a third express frame that was scheduled to leave the CoreSW2 right after express frame 1. It can be seen on the right hand side of the figure that the pacing of frames 1 and 2 remains the same.

Moreover, one must not forget that all the controlled IO devices send their traffic to the controller, thus, the controller link rate has to accommodate all the control traffic. If the

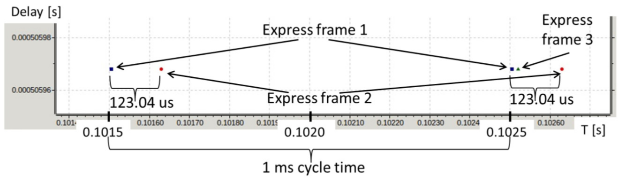

Fig. 9. Inserting a third frame between two express frames (1 Gbps control link)

control link rate is smaller than the core link rate, then it might unnecessarily limit the number of controlled IO devices. In this example with frames of 1500 bytes MTU and 1 ms cycle this upper limit is 8 devices, which is unpractical. An alternative solution would be to lower the frame size: in the same scenario a 150 bytes control frame would allow the control of 81 devices. As a conclusion, the control link rate should be treated as a bottleneck with high impact on the overall capacity of the system during the design of the industrial network.

b) Preemption of small express frames

If the industrial express frames are small, smaller in transmission time than even the preemption delay, then in certain situations the express frames change their order of arrival just because of the preemption delay.

We have set up a small simulation experiment to show this effect, using a cascaded topology with only two express traffic sources and one BE (see Fig. 10). The size of the BE frame is 123 bytes with header, thus it cannot be split into two fragments of 64 bytes, and the preemption delay will take the worst case value, which is the serialization delay of the 123 bytes frame. The link rate at the core is 1 Gbps, all other links are 100 Mbps. The traffic scenario in this network is shown in Fig. 11.

If there is no BE traffic, then express frame 1 is scheduled to

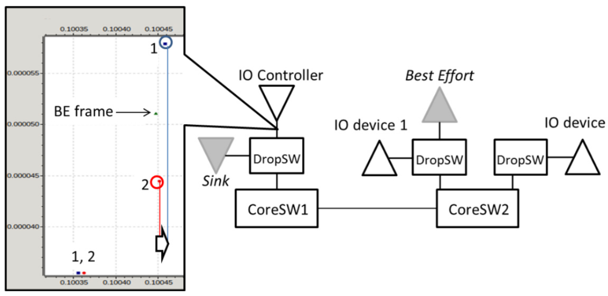

Fig. 10. Evaluating the effect of preemption on very small frames

arrive just before express frame 2 to Core_SW2 (see the upper half of Fig-traffic). When BE traffic is enabled the BE frame is scheduled to arrive to the DropSW on Fig-small just before the express frame 1. When DropSW wants to send express frame 1, the BE frame is already under transmission. Thus, there will be an unsuccessful preemption attempt on the BE frame, and frame 1 gets delayed. Due to this delay express frame 1 will arrive to CoreSW2 later than express frame 2, and thus their order will be swapped, resulting in the output illustrated in the lower half of Fig. 11.

The resulting end-to-end delay values for the case of small express frame size (72 bytes) are plotted in the left hand chart of Fig. 10. The express frames change their order of arrival, and express frame 1 has larger PDV than express frame 2.

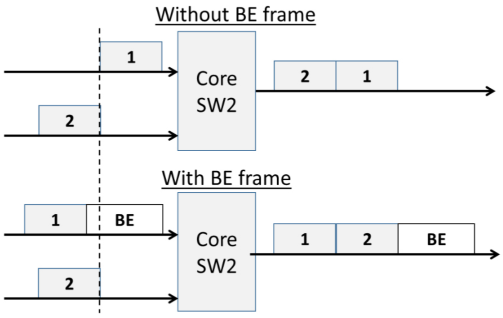

Fig. 11. Preemption scenario with small express frames

We present the resulting PDV values for both cases in Table II. It can be seen that if the express frame size is greater than or equal with 123 bytes, then the PDV is the same for all frames, because there was no change in the order of arrival.

TABLE II
PDV VALUES FOR SMALL EXPRESS FRAME SIZES

Express flow #	72 bytes frame	132 bytes frame
1	19.59 us	9.44 us
2	6.15 us	9.44 us

This shows that using very small frame sizes requires extra care, because PDV can be higher than the serialization delay of a single frame and the intended order of arrival may be changed.

2) Assuming Imprecise Traffic Source Timing

We also investigated the effect of imprecise time synchronization in the network. We simulated this without BE traffic and with modified express frame sizes of 64 bytes on two topologies: cascaded and star (see Fig. 12) with the same arrangement as presented in Section IV.1.

Synchronization errors are modeled at the source by generating a Source Delay Variation (SDV) value. The SDV has a uniform distribution between 0 and maxSDV = 20 ns. This maxSDV value is large enough to cause changes in the order of arrival for express frames from neighboring sources. Note that preemption can protect express traffic against BE traffic only, thus, it will not help in this scenario. Time gating

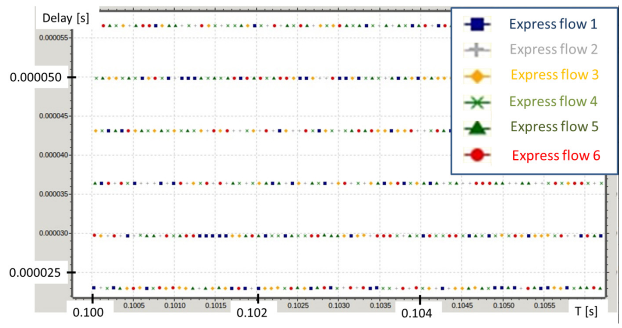

Fig. 12. End-to-end delays of express frames (star topology)

can avoid reordering, but it is also sensitive to synchronization errors. The result is that the PDV equals at least the serialization delay of one frame (6.72 us in this scenario).

Note that PDV also depends on the topology. In case of star topology, in the worst case the affected express frame has to wait for one frame from every other flow, resulting in 5 times larger PDV than the PDV observed in the cascaded topology (see Table III).

TABLE III
PDV VALUES WITH SOURCE DELAY VARIATIONS (MAXSDV = 20 NS)

	Cascaded topology	Star topology
PDV	6.74 us	33.61 us

The conclusion is that the timing of the traffic sources should be designed or shaped by the first switch such that the relative frame arrival time measured at the output link, where they have the first interference, is larger than the maxSDV.

V. CONCLUSION

Due to promises of cheaper operation and converged networking there is much interest in the industry to replace Industrial Ethernet with standard Ethernet solutions. In this paper we have illustrated with simulations that standard Ethernet with TSN features can work as a replacement of Industrial Ethernet solutions. While in PROFINET (and generally in any Industrial Ethernet application) the control and scheduling of the traffic is part of the protocol, once we use standard Ethernet for transport there will be a split between the industrial applications and the transport network.

We have shown that while frame preemption protects the express traffic against the low priority traffic, the interference between express flows must be also resolved, as the total delay can include the serialization delay of other express frames in addition to the preemption delay. With proper time gated queuing configuration the PDV can be eliminated, because this realizes completely scheduled traffic in the network. Standard Ethernet allows using different link speeds in parts of the network, this can be exploited if link speeds are increasing towards the controller. We have also shown that both small and large frame sizes can induce worst case scenarios when using frame preemption.

Therefore, the design of the transport network, including the configuration of the TSN features requires special care. A possible way to handle these issues would be the introduction of a dedicated SDN controller that has an interface to learn the traffic details of the industrial applications.

ACKNOWLEDGMENT

The authors would like to thank the valuable comments received from János Farkas and Balázs Varga from Ericsson and the help of their colleagues István Moldován, József Bíró and Árpád Péter Nagy.

References

[1] H. Kagermann, W. Wahlster, J. Helbig, "Recommendations for implementing the strategic initiative Industrie 4.0", Final report of the Industrie 4.0 Working Group, April 2013.

[2] R.Davies, "Industry 4.0. Digitalisation for productivity and growth", Briefing from EPRS. European Parliamentary Research Service, 2015.

[3] "Digitizing European Industry", European Commission's strategy on Digital Single Market, available from: https://ec.europa.eu/digital-single-market/en/policies/digitising-european-industry, May, 2017.

[4] J. D. Decotignie,"The many faces of industrial Ethernet", IEEE Industrial Electronics Magazine, 3(1), 2009.

[5] P. Danielis et al., "Survey on Real-Time Communication Via Ethernet in Industrial Automation Environments", 2014.

[6] Von Burg, Urs. The triumph of Ethernet: technological communities and the battle for the LAN standard. Stanford University Press, 2001.

[7] M. Fabbi, A. Lerner, "Magic Quadrant for Data Center Networking", Gartner, 2015.

[8] Z. Ghebretensaé, J. Harmatos, K. Gustafsson, "Mobile broadband backhaul network migration from TDM to carrier Ethernet", *IEEE Communications Magazine*, *48*(10), 2010.

[9] Time-Sensitive Networking Task Group description, http://www.ieee802.org/1/pages/tsn.html, 2017.

[10] IEEE 802.1Q-2014 - Bridges and Bridged Networks, http://www.ieee802.org/1/pages/802.1Q.html, 2014.

[11] TTTEch DESwitch Hermes 0/4 Flyer, http://www.tttech.com, February 2017.

[12] Cisco Industrial Ethernet 4000 Series Switches Data Sheet, http://www.cisco.com, May 2017.

[13] Flexible Manufacturing With Time Sensitive Networking (TSN), National Instruments, http://www.ni.com/video/4275/en/, February 2017.

[14] F. de Leow, "Belden demonstrates TSN-ready switches", South African Instrumentation and Control, http://www.instrumentation.co.za/57661N, July 2017.

[15] R. Pigan, M. Metter, "Automating with PROFINET: Industrial Communication Based on Industrial Ethernet", Wiley-VCH, 2008.

[16] ODVA, "The organization that supports network technologies built on the common industrial protocol (cip) - device net, ethernet/ip, componet, and controlnet," Available from: http://www.odva.org, February 2014.

[17] CC-Link Partner Association, "Cc-link ie field brochure," Available from: http://www.cclinkamerica.org, 2013.

[18] SERCOS Automation Bus, Homepage available from: http://www.sercos.com, February 2014.

[19] J. Morse, "World Market for Industrial Ethernet", IMS Research, 2011.

[20] T. Carlsson, "Industrial Ethernet and Wireless are growing fast", Industrial network market shares 2017 according to HMS, February 2017.

[21] IEEE 802.1AS-Rev Timing and Synchronization for Time-Sensitive Applications, http://www.ieee802.org/1/pages/802.1AS-rev.html, March 2017.

[22] IEEE 802.1CB - Frame Replication and Elimination for Reliability, http://www.ieee802.org/1/pages/802.1cb.html, March 2017.

[23] IEEE 802.1Qcc - Stream Reservation Protocol (SRP) Enhancements and Performance Improvements, http://www.ieee802.org/1/pages/802.1cc.html, May 2017.

[24] IEEE 802.1Qbu Frame Preemption, http://www.ieee802.org/1/pages/802.1bu.html October 2015.

[25] IEEE 802.3br Interspersing Express Traffic, http://www.ieee802.org/3/br/, June 2016.

[26] IEEE 802.1Qbv Enhancements for Scheduled Traffic, http://www.ieee802.org/1/pages/802.1bv.html October 2015.

[27] P. Ferrari, A. Flammini, S. Vitturi, "Performance analysis of PROFINET networks", Computer standards & interfaces, 28(4), 369-385, 2006.

[28] A. Verwer, "Introduction to PROFINET", Available from: http://www.profibus.com/, October 2010.

[29] T. Hegr, M. Voznak, M. Kozak, L. Bohac, "Measurement of Switching Latency in High Data Rate Ethernet Networks", Elektronika ir Elektrotechnika, vol. 21(3), 73-78., 2015.

[30] OMNeT++ simulation manual (version 5.0), https://omnetpp.org/doc/omnetpp/manual/, 2017.

[31] INET reference documentation (version 3.4.0), https://omnetpp.org/doc/inet/api-current/neddoc/r.html, 2017.

Csaba Simon is a software engineer and he earned his PhD in 2012 at the Doctoral School of Informatics of the Budapest University of Technology and Economics and currently works as an assistant professor at the Department of Telecommunication and Media Informatics of the same university. His research area includes the topics of Future Internet, 5G networks and cloud systems. He participated in numerous national and international projects in the fields of network resource management, mobility management and smart content delivery. He is a member of Scientific Association for Infocommunications, Hungary (HTE).

Miklós Máté received his MSc (2007) degree in computer science in the field of infocommunication systems at Budapest University of Technology and Economics (BME), Hungary and currently he is preparing to defend his PhD dissertation at the same University. He is a research engineer in the High- Speed Networks Laboratory at the Department of Telecommunication and Media Informatics, BME. His research interests include intelligent transportation systems and distributed networks.

Markosz Maliosz received his MSc (1998) and PhD (2006) degrees in computer science in the field of infocommunication systems at Budapest University of Technology and Economics (BME), Hungary. He is an associate professor in the High- Speed Networks Laboratory at the Department of Telecommunication and Media Informatics, BME. His research interests include virtual, cloud and sensor networking along with optimization techniques. He participated in numerous national and international projects in the fields of network resource management, multimedia and smart content delivery. He is a member of Scientific Association for Infocommunications, Hungary (HTE).

Norbert Bella received his BSc degree in the field of infocommunication systems at Budapest University of Technology and Economics (BME), Hungary. He participated in the research project at the same institution evaluating several use cases of the Time-Sensitive Networking.