Next Generation Real-Time Networks Based on IT Technologies

Wilfried Steiner*, Pablo Gutiérrez Peón*, Marina Gutiérrez*, Ayhan Mehmed*,
Guillermo Rodriguez-Navas†, Elena Lisova†, Francisco Pozo†
*TTTech Computertechnik AG, Vienna, Austria
†Mälardalen University, Västerås, Sweden

***Abstract*—Ethernet-based networks have found their way into industrial communication more than a decade ago. However, while industry and academia developed Ethernet variants to also meet real-time and fault-tolerant requirements, recent standardization efforts within the IEEE 802 will broadly bring standard IT switched Ethernet in future industrial communication networks. As first standards of IEEE 802.1 time-sensitive networking (TSN) are becoming published at the time of this writing, we review these standards and formulate further research challenges that still go beyond current standard developments. Furthermore, we report on recent research results from the RetNet project that target these research challenges.**

I. Introduction

Today we live through the beginning of a technological revolution in the area of dependable and secure systems driven by an accelerated integration of information technology (IT) and operations technology (OT). IT encompasses, e.g., advanced networking technologies, like Ethernet, TCP/IP, and higher layered networks, software-defined networking (SDN), or virtualization technologies. OT, on the other hand, encompasses technology typically associated with embedded systems and cyber-physical systems. Sometimes IT and OT already overlap today and various industry trends indicate a much stronger need for integration of IT and OT in the near future. Examples of these IT/OT integration trends are: Industrie 4.0, Internet of Things, intelligent transportation systems, smart grid, smart city, and other smart developments. This integration will be one of the main motors that contributes to the success of these developments.

IT/OT integration is also at the core of the definition of new standards for network communication. Specifically, the IEEE 802, which traditionally defined protocols only for information communication, has been working on the definition of new IT-based standards for embedded system communication. In particular we refer to the developments that origin in the IEEE 802.1 AVB (Audio/Video Bridging) and IEEE 802.1 TSN (Time-Sensitive Networking) standardization groups, and that are applicable into fields, such as factory automation, industrial automation, substation control, process control or similar.

The core of these novel standards, Switched Ethernet, is a very successful technology in the IT domain, which for some years has been regarded as the best candidate to replace (or at least complement) fieldbus technology in industry [1]. Indeed, a multitude of Ethernet variants has been developed by industry for OT use, as for example: PROFINET, Ethernet Powerlink, and EtherCAT. In parallel, oriented basic academic research in the area of factory communication proposed real-time and fault-tolerant extensions to Ethernet (e.g., [2] [3]). For defining new functionalities within the native Ethernet standardization body, IEEE 802.1 TSN can, thus, build on both: (a) a solid foundation of research results paired with OT industrial expertise and (b) the existing strengths of IT Switched Ethernet, like high speed, low component cost, dynamic topology management, routing algorithms, transport data protocols and internal switch operation. In Section II we review the current state of the standardization of IEEE 802.1 AVB and IEEE 802.1 TSN, with special emphasis on the recently-added features. It is shown that an overarching characteristic of these protocols is the consideration of *time* as a "first-class citizen"; something very different from the usual IT perspective, in which *performance* is paramount.

978-1-5090-1314-2/16/$31.00 ©2016 IEEE

Despite the advancement of the IEEE 802.1 standards, there are still a number of open challenges for the development of time-sensitive large-scale networks for complex cyberphysical systems. These aspects are being addressed in the EU FP7 project RetNet, and include integration of wired and wireless Switched Ethernet technology, large-scale offline scheduling of time-triggered traffic, runtime auto-configuration and adaptation for time-sensitive networks and security of time-sensitive networks. In Section III we discuss the main results already achieved in RetNet as well as the next steps in the project. Finally, Section IV summarizes and concludes the paper.

II. IT Networks for Factory Automation

A. Basic IEEE 802.1 Concepts

We will refer to IT Ethernet for OT usage as Deterministic Ethernet in this paper. Such a Deterministic Ethernet, builds on switched Ethernet, i.e., end nodes are not directly connected to each other, but are connected to switches that forward the messages from a sending end node, either to the receiving end node, or to another switch in closer proximity to the receiving node. Hence, switches may be connected to each other forming an arbitrary physical network topology. On top of the physical network topology one or many logical topologies can be imposed by selecting a set of the underlying physical connections. Logical topologies are either manually configured, or evolve dynamically by the execution of network protocols, such as the Spanning Tree protocol or the Shortest Path Bridging (SPB) protocol.

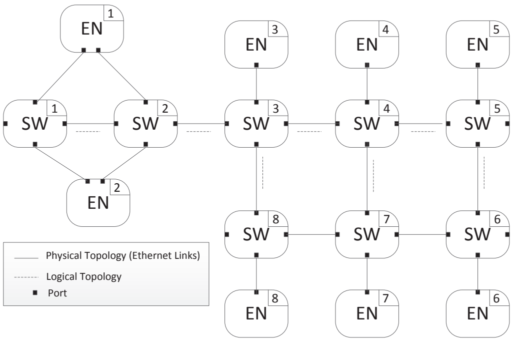
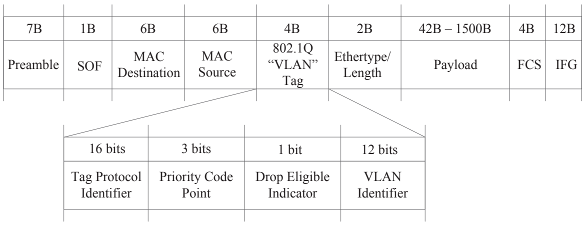


Fig. 1. A network example

An example of an Ethernet network consisting of eight end nodes (EN 1-8) and eight switches (SW 1-8) is depicted in Figure 1. The solid lines indicate physical Ethernet connections that interlink the end nodes to the switches and the switches to each other. The dashed lines depict a logical topology forming a spanning tree, i.e., the dashed lines form a logical topology in which each end node and switch may reach any other end node and switch via exactly one route through the network. In this example, the physical Ethernet links connecting switches SW 6 to SW7 and SW 7 to SW 8 are excluded from the logical topology. Furthermore, in this example a single spanning tree is depicted. More recently, Ethernet switches implement shortest path bridging protocols such as IS-IS SPB (Intermediate Station to Intermediate Station Shortest Path Bridging - IEEE 802.1aq). While in the example presented in Figure 1 the shortest path from switch SW 1 to any other switch and node is presented and used for all communication in the system, IS-IS SPB would install a dedicated shortest path for each switch in the network to each other switch, i.e., IS-IS SPB configures multiple spanning trees to form the logical topology of the network. For example, the shortest path for SW 7 to communicate with switch SW 8 would indeed use the direct link between these two switches (instead of the communication through SW 4, SW 3, to SW 8).

Once the logical topology is established the network transports data in form of Ethernet messages[1] only on those Ethernet links that belong to the logical topology. The Ethernet frame format is depicted in Figure 2. It consists of a frame preamble and a start of frame pattern (SOF) followed by addressing information about the message's destination (MAC Destination) and source (MAC Source). Typically, the source and destination addresses are used by the Ethernet switches to determine the communication path of the respective message. An optional "VLAN" tag (introduced by IEEE 802.1Q) can be used and we discuss its functionality later in this paper. The Ethertype field defines which protocol (if any) are being transported in this frame. Alternatively, for small messages the Ethertype field can be used to indicate the frame's length. Payload holds the actual data (plus optional higher layer metadata). Finally, the Ethernet frame concludes with a frame check sequence (FCS) - a thirty-two bit long cyclic redundancy check. In between two Ethernet frames a minimum inter frame gap (IFG) needs to be respected.



7B	1B	6B	6B	4B	2B	42B – 1500B	4B	12B
Preamble	SOF	MAC Destination	MAC Source	802.1Q "VLAN" Tag	Ethertype/ Length	Payload	FCS	IFG

16 bits	3 bits	1 bit	12 bits
Tag Protocol Identifier	Priority Code Point	Drop Eligible Indicator	VLAN Identifier

Fig. 2. Ethernet frame format

With respect to the OSI network layers, IEEE 802.3 defines Ethernet layer 1 (physical layer - PHY) and layer 2 (media access layer - MAC). Complementary to the IEEE 802.3, the IEEE 802.1 working group standardizes protocols that use Ethernet messages as well as functions and mechanisms to handle messages in the switches and end nodes. As in this paper we are concerned with the real-time, safety, and availability enhancements for Ethernet we focus mostly on the IEEE 802.1 developments.

B. Detailed Switch Operation

Ethernet switches (or "bridges" as defined in the IEEE 802.1 standards), differentiate between the control-plane and the data-plane. The control-plane is concerned with all functionality that enables the data-plane to operate. Example control-plane functions are therefore: the establishment of logical topologies, network configuration in general, or certain form of diagnostics.

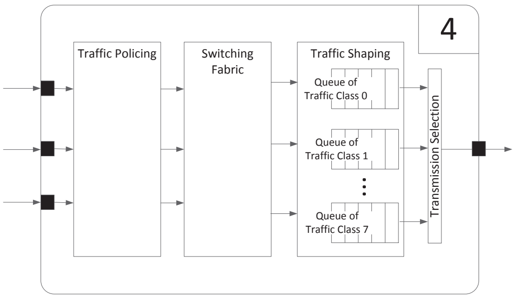


Fig. 3. Example of the data-plane in a switch

The data-plane is typically the time-critical aspect of the switch and is in charge of forwarding messages from its inputs to the outputs. The data-plane of an Ethernet switch is sketched in Figure 3. We assume that this example switch represents switch SW 4 in the illustrative network in Figure 1. The black rectangles represent physical ports of the switch and in this scenario we assume that SW 4 receives messages on three

[1] We use the terms "message" and "frame" synonymously.

ports (on the left) and forwards these messages on an outgoing port (on the right). As depicted, while the switch processes the frames from input to output, three main functions can be distinguished: traffic policing, the actual switch fabric, and traffic shaping.

Traditionally, (i.e., before IEEE 802.1 TSN has been started) a purely standards-based Ethernet bridge has been using only very few parameters in this forwarding process:

- The output ports for an incoming message have only been defined by the MAC Destination address of the incoming message.
- The VLAN tag allowed to restrict the set of eligible ports to classes of messages.
- Only the priority code point in the VLAN tag has defined the priority queue into which a message would be inserted to at the outgoing port.

Various vendor-specific extensions have been designed to overcome technological shortcomings of this simple approach. Especially in the area of industrial communication and factory automation a multitude of Ethernet-based variants have been developed that extended this simple forwarding principle in order to achieve better real-time, safety, and/or availability properties as often required in OT use cases. TSN is currently extending the parameters in the forwarding process, e.g.:

- Switches may use a network-wide (or sub-network-wide) synchronized global time in the traffic policing and traffic shaping process.
- Switches may identify specific flows of messages (i.e., messages with the same identifier) in the forwarding process. This is in contrast to the class-based treatment of messages based on the VLAN priority code point.

In addition to these parameters TSN defines new forwarding functions, which we discuss next.

1) Traffic Policing: Traffic policing is the process of checking whether incoming messages adhere to certain acceptance rules defined in the switch, or not. Such rules can be both, in the value domain and in the time domain. Checks in the value domain evaluate if specific bits or bit combinations in the incoming message are set in adherence with the acceptance rules. Checks in the time domain, on the other hand, evaluate whether the receive points in time of messages are correct or not. These checks in time can be further divided into synchronized checks and unsynchronized checks. Synchronized time checks will evaluate whether the point of time of reception of the respective message is correct with respect to a synchronized global time, while the unsynchronized checks evaluate whether the temporal distance between successive messages of the same flow or message class is in defined bounds (typically only a lower bound on temporal distance is checked). When traffic policing finds an incoming message to violate an acceptance rule the message is either discarded or re-prioritized to a lower priority level. In this case, the drop eligible indicator flag in the VLAN tag may also be set to indicate that the message may be dropped in case of congestion at the outgoing port.

The ongoing IEEE 802.1Qci project of TSN is currently standardizing traffic policing and filtering methods.

2) Traffic Shaping: Traffic shaping is the actual scheduling process executed at the outgoing port. As depicted in Figure 3, typically multiple queues feed into a single physical output port. Hence, when multiple queues at a port are not empty, a scheduling decision needs to define the queue from which the next message will be selected. The most simple approach is strict priority in which messages from higher priority queues are scheduled for transmission before messages from lower queues.

AVB has standardized the credit-based shaper that defines a local variable, the *credit*, per queue (that has the shaper activated). This value of the credit is increasing with a rate of “idleSlope” when there are messages in the respective queue ready but are not being served. The value of the credit is decreased with a rate of “sendSlope” when the respective queue is being served. Messages of a queue are only scheduled for transmission if the respective credit is positive. The credit-based shaper ensures that even high-priority queues will not monopolize the outgoing port (as it would be possible with strict priority scheduling).

TSN has standardized the time-aware shaper (IEEE 802.1Qbv) that takes into account a synchronized network-wide time for its scheduling decisions. In particular, each queue can be in an active or inactive state and the switch will select messages from those queues that are in the active state. The state of each queue can change between active and inactive in accordance with an offline defined communication schedule.

TSN is currently also about to finalize the cyclic-queuing and forwarding shaper in IEEE 802.1Qch. This shaper extends the time-based shaper IEEE 802.1Qbv. It takes the synchronized time also into account when deciding which queue the incoming message is assigned to. Thus, a message with a certain identifier can be assigned to different queues depending on which time it has been received by the switch.

Most recently TSN is starting a project to standardize an asynchronous traffic shaper (IEEE 802.1Qcr) that aims for a similar level of time-critical communication as the time-aware shapers, but without the reliance on synchronized time.

3) Switch Fabric: The switch fabric is in charge of the actual data transmission between input and output. While in the past cross-bar switches with local memories on input ports an output ports were common, today’s switches often implement a centralized memory structure as the switch fabric.

Switches can operate in store-and-forward as well as in cut-through mode. Store-and-forward means that a message must be completely received by the switch before the switch can start its transmission on the outgoing ports. Cut-through, refers to the complementing concept – message transmission can be started even if the message has not been completely received by the switch. It is an ongoing debate whether or not cut-through switches actually are covered by the IEEE 802.1 and 802.3 standards, or not. However, multiple vendors implement cut-through functionality. Furthermore, TSN IEEE 802.1Qcc standardizes configuration parameters for cut-through, thereby arguing in favor of a standardized cut-through behavior.

In addition to store-and-forward and cut-through, TSN has also standardized a message preemption service in IEEE 802.1Qbu and 802.3br. This allows a high-priority Ethernet frame to interrupt a lower-priority Ethernet frame. Once the high-priority frame transmission or transmissions are completed the lower-priority frame transmission can resume.

C. Redundancy Management

TSN standardizes redundancy management in two main standards. First, IEEE 802.1Qca standardizes automatic ways to configure multiple routes through a network. To do so, the standard extends the capabilities of the IS-IS protocol. Furthermore, the IEEE 802.1CB standard defines how to identify and to merge redundant copies of the same message at intermediate points in the network as well as at the final receivers. Due to space limitations of the paper these techniques will not be described in detail. We simply note that the IEEE 802.1CB techniques are similar to techniques used in the High-availability Seamless Redundancy protocol (HSR) and the Parallel Redundancy Protocol (PRP). Indeed, at the time of this writing the IEEE 802.1 is aiming to establish a formal liaison with the respective standardization groups in the IEC standardization group.

D. Network Management

Traditionally, plug-and-play has been a core element of IEEE 802.1 networks – a user would simply connect various devices to each other and the IEEE 802.1 protocols would automatically establish connectivity. SRP (stream reservation protocol) is an example of such a protocol that has been standardized within the IEEE 802.1 AVB task group (in 2011). With SRP, end stations can advertise streams to the network and other end stations can register to these advertised streams. Functionality in the bridges ensures that the overall number of streams is below a given threshold such that defined maximum end-to-end latencies of the streams can be guaranteed.

While SRP is sufficient for typical use cases in consumer electronics and even professional audio/video processing, it has shortcomings for applications with more demanding requirements, such as industrial control or automotive control applications. The IEEE 802.1 TSN task group has therefore identified the need to improve the configuration and management capabilities and IEEE 802.1Qcc is currently standardizing such improvements. In the following we discuss the main concepts most likely to become standardized within the next twelve months of the time of this writing.

The configuration data of a TSN network is represented as a set of managed objects formulated as YANG models [4] and IEEE 802.1Qcc differentiates two concepts on how the managed objects may be modified: user/network interface (UNI) and remote management. Remote management addresses the TSN configuration from a system's administrator-like perspective. It allows a remote entity, like an end station (probably a laptop computer), to read/write/modify managed objects residing locally in a given bridge in the network. Hence, remote management allows a fine-grained access to a bridge's configuration. IEEE 802.1Qcc does not define, nor require, a specific remote management protocol. However, such protocols are widely known and used. Examples of remote management protocols are SNMP or NETCONF. The UNI, on the other hand, addresses the TSN configuration from a user-network interaction perspective. Thus, while remote management allows access to all managed objects in all bridges, the user will typically have only limited access to the managed objects through a UNI. In an extreme case, the UNI may only allow the user to formulate specific requirements on the network, e.g., a target network latency for a given stream, and the network may only respond to the user through the UNI whether the requirements have been met, or not.

Based on the differentiation of the UNI from remote management, IEEE 802.1Qcc defines three models of configuration: fully distributed model, centralized network / distributed user model, and the fully centralized model. We discuss the three models in detail next.

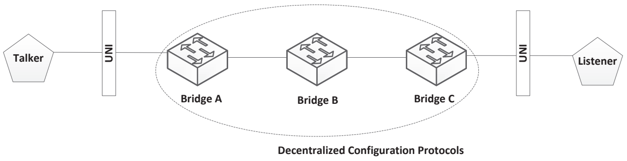


Fig. 4. The fully distributed configuration model

In the fully distributed model the end stations (talkers and listeners) are exposed to the network through the UNI as depicted in Figure 4. Inside the network (communication between bridges A-C), the bridges coordinate themselves via decentralized configuration protocols. SRP is an implementation of this model: talkers advertise their capabilities to the network and listeners register to advertised streams. The managed objects inside the bridges are largely hidden from the end stations and are mostly local information to each individual bridge.

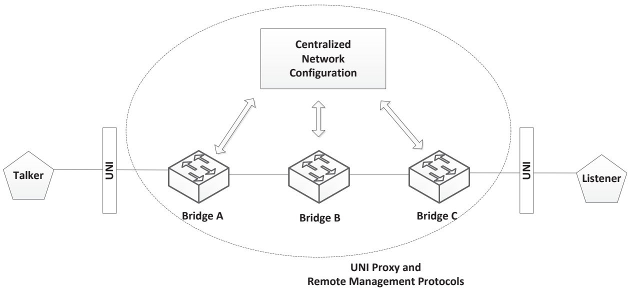


Fig. 5. The centralized-network distributed-user configuration model

The centralized-network distributed user configuration model is depicted in Figure 5. Again, the talkers and listeners expose their communication capabilities and needs to the network through the UNI. However, in contrast to the fully distributed model above, here, the bridges do not coordinate their operation themselves. Instead, the edge bridges (i.e., bridges that connect to end stations like Bridge A and Bridge B) communicate the information received from the users (through the UNI) to a central network configuration entity (CNC). The CNC will typically have a much broader and more complete view on the network and therefore can produce more efficient network configurations than achieved by decentralized protocols. For example, the CNC can use remote management protocols to gather information from the bridges regarding their performance characteristics. Likewise, once the CNC has produced a new network configuration, it will use remote management protocols to download this configuration to the bridges.

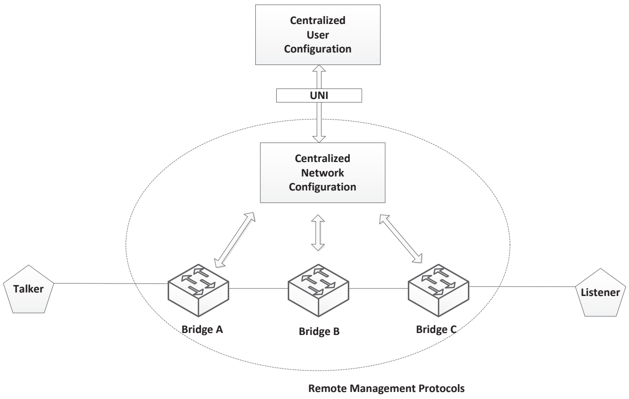

Centralized User Configuration
UNI
Centralized Network Configuration
Talker
Listener
Bridge A
Bridge B
Bridge C
Remote Management Protocols

Fig. 6. The fully-centralized configuration model

The fully centralized model is depicted in Figure 6. This model assumes the presence of a centralized user configuration (CUC) entity. The CUC is assumed to replace the need for UNI between each end station and the network. Instead, the CUC has complete knowledge of the end stations' communication needs. The CUC can acquire such detailed knowledge through direct user interaction, e.g., the system administrator registers a new end station and its capabilities and needs with the CUC. Alternatively or complementary, other protocols, outside the scope of IEEE 802.1Qcc (and probably outside IEEE 802.1 entirely), can be executed on the network to establish the required knowledge in the CUC. Conceptually, the CUC can then interact with the CNC through a UNI (or an UNI-like interface) to inform the CNC of the communication needs between the end stations. Thus, in contrast to the previous configuration model, bridges in the fully centralized model do not need to operate as proxies for the end stations.

III. RetNet Project: Goal and Status

RetNet stands for the *European Industrial Doctorate Programme on Future Real-Time Networks*. The main scientific goal of this EU FP7 project is to advance the state of the art and practice in predictable real-time and Internet networking technologies. This goal has been structured in four different *challenges*, which will be presented along this section.

A. Wireless Challenge

Industrial communications have traditionally relied on wired fieldbus systems, that are able to provide real-time guarantees for applications with low throughput requirements [5]. In contrast, the use of wireless communication systems in factories have remained marginal until recently. The new possibilities started to be explored when the industry became aware of the many benefits of wireless communication, and realized they could be a good complement to the existing and already mature wired networks. Benefits of reduced wiring include easier deployment, decreasing costs, and easier operation in equipment with moving components. Wireless communication also exhibits some limitations: it operates only as half-duplex, which implies reduced bandwidth unless more than one channel and antenna are used; transmissions on the same frequency band are broadcasted within the range, so concurrent transmissions cannot take place; and the wireless signal is subject to drastic changes (mainly caused by multipath fading, shadowing and interference) that can compromise the quality of the communication. Diversity techniques based on sending the same information through channels with different characteristics may counteract multipath fading and shadowing, increasing communication quality [6]. Interference are caused by other wireless devices transmitting simultaneously and need to be handled by other means, like proper node coordination.

The goal of the RetNet project is to provide wireless IT technologies for the OT in the context of factory automation, while maintaining the benefits of such technologies: high data rate, low cost and ease of use. RetNet also addresses problems associated to so-called *hybrid* networks, in which data can seemingly be transmitted through wired and wireless communication systems. This case is particularly challenging because the network is no longer homogeneous and has to cope with e.g., different capacities or speeds between the wired and wireless segments.

RetNet also focuses on the design of proper MAC protocols to prevent devices from performing uncoordinated transmissions within the same network, and therefore eliminate the possibility of interference. In this context, IEEE 802.11 for wireless local area networks and IEEE 802.15.4 for wireless personal area networks are prominent examples of widely used standards in the IT and OT fields respectively. However, they are mainly based on the use of the carrier sense multiple access (CSMA) MAC protocol, which provides random access to the medium and is consequently susceptible to interference. A different approach is taken by IsoMAC [7] and WirelessHART [8], that provide time-critical communication based on IEEE 802.11 and IEEE 802.15.4 respectively. Both technologies employ time division multiple access (TDMA) on top of CSMA. TDMA works by dividing the medium access into time-slots that can be unambiguously assigned to a sender, so that the access to the medium is guaranteed. The schedule that assigns time-slot to senders must be known by all network participants. Additionally, the instants when a time-slot starts and ends must be consistent between the network stakeholders, implying that some sort of synchronization protocol is needed.

The standard amendments under TSN, discussed in Section II, can also be considered in order to achieve real-time communication guarantees in wireless. The time and synchronization aspects included in IEEE 802.1AS can be applied, in the same way it has been done before with other synchronization protocols like IEEE 1588 [9]. Under the assumption of having clock synchronization available to wired and wireless alike, MAC protocols based on TDMA can be adopted. The enhancements for scheduled traffic considered in IEEE 802.1Qbv can also be translated to wireless. The time-aware shaper relies on a schedule that enables or disables the transmissions coming from different queues. In wireless, the schedule should guarantee that the queues are not enabled concurrently for devices under the same range.

Since the MAC layer is a key component for providing real-time guarantees, the RetNet project has so far focused on providing a wireless MAC protocol suitable for working in hybrid networks with heterogeneous traffic, i.e. time-critical and best-effort traffic. Our proposal [10] is that time-critical traffic is scheduled offline for both wired and wireless segments. The size of a slot is calculated after considering the largest possible data transmission plus protocol overhead. In hybrid networks, messages may pass through media with different transmission speeds, and that would involve different slot sizes that have to be accounted for by the scheduler, or otherwise the slot size would have to be as large as required for accommodating a transmission in the slowest transmission medium. After allocation of time-slots for time-critical traffic, the remaining time, if any, can be used by best-effort traffic. We have evaluated several MAC protocol variants that allow to balance off-line scheduling versus on-line flexibility for best-effort traffic [11].

B. Scheduling of Large Systems

A schedule can be seen as a consistent agreement between all the end systems and switches that indicates, during all system operation, the points of time in which each time-triggered message is transmitted and over which link. The importance of using a global schedule of messages, based on a system-wide global clock, has been acknowledged by many standards for deterministic communication, as highlighted in Section II. The schedule is a fundamental element of more advanced mechanisms at the data-plane, like traffic policing and traffic shaping. In wireless communication, it improves coordination among nodes, reducing interference.

Composing a global schedule is known to be NP-complete with the size and the number of messages in the network. Multiple techniques and tools have been used for synthesizing schedules small to medium size networks. Steiner [12] translated the scheduling problem into integer linear constraints formulas that are fed into a Satisfiability Modulo Theories (SMT) solver which determines the satisfiability of such formulas and returns a communication schedule (when the constraints are satisfiable). Moreover, in order to be able to solve larger networks, an incremental strategy to synthesize the schedule by smalls steps was designed. Craciunas et al. [13] developed communication and task co-synthesis schedules using an SMT-based approach. Alternatively to SMT solvers, also Mixed Integer Programming (MIP) solvers have been proposed, which allow optimization in different parameters. Combination of different solver techniques are also used, e.g., an Answer Set Programming (ASP) solver interacts with the SMT solver in order to synthesize the schedule [14]. Not only constraint solvers are used, specialized search tools, such as meta-heuristics are also able to construct time-triggered schedules: Tamas et al. use tabu-search to synthesize schedules of time-triggered messages taking into account unsynchronized (aperiodical) messages [15].

These approaches are able to synthesize medium-scale networks up to a thousand of messages in less than an hour. However, upcoming systems requiring OT, will consist of much larger networks and up to hundreds thousand frames, whereby all approaches mentioned present serious scalability issues. RetNet has researched divide and conquer techniques for managing complexity and reducing the synthesis time. The approach presented in [16] that divides the global schedule in small schedules called segments. Each segment consists of a fraction of the global schedule in which messages are allocated using some scheduling approach. However, complications arise when constraints exist between two frames that belong to different segments, e.g., application constraints between frames, in which a frame transmission must be scheduled a defined amount of time before the transmission of another frame. A segment handler is introduced to decide which frames are going to be scheduled in each segment and to add new inter-segment constraints to the segment scheduling problem to satisfy constraints between frames of different segments. This decomposition approach is able to synthesize up to a hundred-thousand frames in less than one hour. However, further enhancements need to be done in order to apply different types of inter-segments constraints in a single schedule.

C. Autonomous Configuration and Optimization

The introduction of the time-aware shaper in TSN turned network configuration a more complicated task. Concretely, as it has been explained in Section III-B, synthesizing a time-triggered schedule is not a trivial problem. But obtaining the parameters needed to produce the schedule also adds to this complexity, since it requires complete knowledge of the network and the applications using it [12]. This makes reconfiguration of such networks costly both in terms of time and engineering effort, and requires some down time of the system [17]. For this reason, flexible reconfiguration solutions have been proposed over a decade ago, e.g., by Pedreiras et al. [2]. RetNet has introduced the concept of *configuration agent*, for configuration of IT networks for OT use [18]. This agent is an autonomous entity that learns the characteristics of the network through continuous monitoring, facilitating online configuration and re-configuration of time-triggered networks.

The four elements that compose the configuration agent can be seen in Figure 7. The *monitor* observes the network

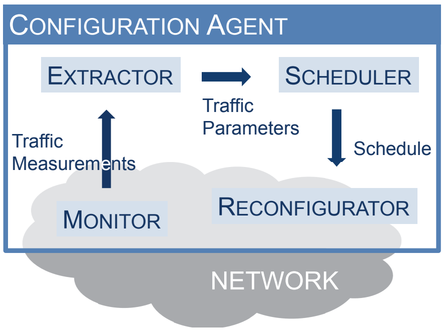


Fig. 7. Configuration agent overview [19]

and gathers traffic measurements, such as the ingress and egress times of messages. From those traffic measurements the *extractor* distills the traffic parameters that the *scheduler* needs to produce a new schedule for the networkk, like the period of the messages, their priorities, and precedence relationships between them. Finally, the *reconfigurator* is in charge of updating the network configuration such that the new schedule is followed.

The duple formed by the monitor and the extractor conform the learning phase of this re-configuration approach and it has been evaluated by means of simulation in [19]. There the monitor tracks the arrival times of messages to the switches. The results showed that the period of the messages, one of the most critical message parameters for the time-aware shaper, can be learned with high accuracy, even in scenarios with high utilization.

The configuration agent fits very well with the software defined networking (SDN) approach [20]. This emerging paradigm defines the network behavior via software tools thus increasing the flexibility of the network. The three main characteristic of SDN are the separation of the control plane from the data plane, the programability of the network and a controller that has a centralized view of the network and can control network devices.

The elements of the configuration agent can be integrated in an SDN architecture. Whereas the monitors need to be placed in the switches, the scheduler and reconfigurator need a centralized view and therefore can be part of the SDN network controller. The functionality of the extractor could be fully implemented also in the network controller or a pre-extraction can be done locally in the switches. This latter option of the two-step extraction can alleviate the traffic workload introduced by configuration agent. The SDN-like architecture for the configuration agent is well in line with the centralized models proposed in IEEE 802.1Qcc and the use of NETCONF as the management protocol and YANG language to model the managed objects.

D. Security aspects of time-sensitive networks

From a security perspective, the adoption of IT technologies for OT is a two-edge sword. On the one hand, such a mix of IT and OT technologies allows industrial applications to keep up with the growing level of network complexity and sub-networks interconnections. On the other hand, the introduction of IT technologies in OT environments significantly increases the security risk and novel security frameworks need to be developed to meet industrial requirements. Among these, timeliness, availability, reliability and heterogeneity can be defined as the main requirements and related to the vast majority of networks, especially considering safety-critical applications [21]. RetNet focuses on these requirements and investigates how they are supported by current solutions and how they can be improved.

Security solutions usually imply additional communication overhead in networks, negatively influencing the support of real-time requirements. Heterogeneity support means providing security solutions for mixed wired and wireless network with different types of traffic, e.g., cyclic, aperiodical and sporadic traffic with no guarantee about message delivery. For security, it means having different policies for different traffic classes and flexible solutions that can be sufficiently reconfigured depending on message priority. Wired networks are usually considered as closed and static, so they often lack dedicated security protocol layers, but are simultaneously considered safe enough under assumption of closed environment.

Adoption of wireless solutions for OT applications is just starting. Because of the open nature of communication, wireless links security aspects are usually considered from the beginning of the development process. However, still such solutions are traditionally considered as not reliable enough. One of the target of the security research in an industrial context is to bring safety and security to acceptable levels for networks containing both wired and wireless solutions. Availability as a security objective means that the network needs to be robust enough to sustain malicious attacks, e.g., clock synchronization should continue working in presence of a malicious adversary trying to disrupt it. Reliability requires that services provided by the system should not deviate from what is expected to be delivered even operating in malicious environment.

Besides the main requirements there is a set of specific ones, like memory and performance limitations, energy efficiency or lower delay. A concrete set of requirements tightly depends on the use case and threat model developed for the application, as RetNet has investigated in [22].

These challenges coming along with providing security services complying with industrial requirements and the evaluation of existing security solutions [23] show the need for a security framework capable of securing main system assets along with relevant requirements support. The development of such a framework is a part of the RetNet project. For framework development, an iterative approach has been chosen. First, possible system assets are evaluated and analyzed from a security point of view. Each iteration brings security solutions for current considered system asset protection. At the current stage of research, mechanisms of clock synchronization protection are considered [24]. Clock synchronization is an essential asset for most real-time networks. To be able to cooperate and follow the schedule, network participants need to share the same notion of time, i.e., they need to be synchronized. One popular standard for clock synchronization is IEEE 1588. This standard has some guidelines for security services in Annex K, however, these measures cannot sustain simple combination of ARP poisoning attack with selective delay attack. This example shows the vulnerability of one of the main assets of many industrial networks. As a possible solution network distributed monitoring was proposed. The possible interactions between adversary targeting clock synchronization breaching and the monitor protecting the network are evaluated from a game theory point of view.

Bringing together IT and OT seems to be a promising trend, as it allows to gain advantages from IT technologies in OT applications. However, these new solutions incorporating both should have a security level responding to nowadays market demands. Therefore, existing security solutions need to be re-evaluated and new solutions need to be developed to complete the framework capable of coping with specific challenges in the area. This is an overall goal of the security research within the RetNet project.

IV. Conclusion

Today, various standards that define IT communication systems (e.g., IEEE 802) are being extended to make them applicable for OT use and large IT semiconductor providers are in the process to adopt their solutions towards these extensions. Thus, in the future, low-cost Ethernet equipment will become available with the goal to consolidate the multitude of today's real-time Ethernet solutions. In this paper we have outlined these developments. However, these protocols and products will not instantly solve all industrial communication issues. As we have discussed in this paper various research questions remain open and novel research problems evolve. Some of these research aspects are in the areas of wireless communication, offline configuration of large-scale schedules, auto-configuration and traffic adaption, and security. The RetNet project is studying these aspects and in this paper we have also reported first project results.

Acknowledgments

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement 607727.

References

[1] B. Galloway and G. P. Hancke, "Introduction to industrial control networks," *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 2, pp. 860–880, 2013.

[2] P. Pedreiras, L. Almeida, and P. Gai, "The ftt-ethernet protocol: Merging flexibility, timeliness and efficiency," in *null*. IEEE, 2002, p. 152.

[3] H. Kopetz, A. Ademaj, P. Grillinger, and K. Steinhammer, "The time-triggered ethernet (tte) design," in *Object-Oriented Real-Time Distributed Computing, 2005. ISORC 2005. Eighth IEEE International Symposium on*. IEEE, 2005, pp. 22–33.

[4] M. Bjorklund, "Rfc 6020: Yang-a data modeling language for the network configuration protocol," 2010.

[5] P. Pleinevaux and J.-D. Decotignie, "Time critical communication networks: Field buses," *IEEE Network*, vol. 2, no. 3, pp. 55–63, 1988.

[6] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[7] H. Trsek and J. Jasperneite, "An isochronous medium access for real-time wireless communications in industrial automation systems-A use case for wireless clock synchronization," in *IEEE Symposium on Precision Clock Synchronization for Measurement, Control and Communication (ISPCS)*, Munich, Germany, 2011, pp. 81–86.

[8] J. Song, S. Han, A. K. Mok, D. Chen, M. Lucas, and M. Nixon, "WirelessHART: Applying wireless technology in real-time industrial process control," in *IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS)*, St. Louis, USA, 2008, pp. 377–386.

[9] T. Cooklev, J. C. Eidson, and A. Pakdaman, "An implementation of IEEE 1588 over IEEE 802.11b for synchronization of wireless local area network nodes," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 5, pp. 1632–1639, 2007.

[10] P. Gutiérrez Peón, E. Uhlemann, W. Steiner, and M. Björkman, "A Wireless MAC Method with Support for Heterogeneous Data Traffic," in *41st Annual Conference of the IEEE Industrial Electronics Society (IECON)*, Yokohama, Japan, 2015.

[11] P. Gutiérrez Peón, E. Uhlemann, W. Steiner, and M. Björkman, "Medium Access Control for Wireless Networks with Mixed Criticality Real-Time Requirements," in *42nd Annual Conference of the IEEE Industrial Electronics Society (IECON), to be published*.

[12] W. Steiner, "An evaluation of SMT-based schedule synthesis for time-triggered multi-hop networks," in *Real-Time Systems Symposium (RTSS), 2010 IEEE 31st*. IEEE, 2010, pp. 375–384.

[13] S. S. Craciunas and R. S. Oliver, "SMT-based task-and network-level static schedule generation for time-triggered networked systems," in *Proceedings of the 22nd International Conference on Real-Time Networks and Systems*. ACM, 2014, p. 45.

[14] A. Biewer, J. Gladigau, and C. Haubelt, "Towards Tight Interaction of ASP and SMT Solving for System-Level Decision Making," in *Architecture of Computing Systems (ARCS), 2014 27th International Conference on*. VDE, 2014, pp. 1–7.

[15] D. Tamas-Selicean, P. Pop, and W. Steiner, "Synthesis of communication schedules for TTEthernet-based mixed-criticality systems," in *Proceedings of the eighth IEEE/ACM/IFIP international conference on Hardware/software codesign and system synthesis*. ACM, 2012, pp. 473–482.

[16] F. Pozo, W. Steiner, G. Rodriguez-Navas, and H. Hansson, "A Decomposition Approach for SMT-based Schedule Synthesis for Time-Triggered Networks," in *Emerging Technologies & Factory Automation (ETFA), 2015 IEEE 20th Conference on*. IEEE, 2015, pp. 1–8.

[17] L. Dürkop, , J. Jasperneite, and A. Fay, "An Analysis of Real-Time Ethernets With Regard to Their Automatic Configuration," in *11th IEEE World Conference on Factory Communications Systems (WFCS)*, 2015.

[18] M. Gutiérrez, W. Steiner, R. Dobrin, and S. Punnekkat, "A configuration agent based on the time-triggered paradigm for real-time networks," in *11th IEEE WFCS*, 2015.

[19] M. Gutiérrez, W. Steiner, R. Dobrin, and S. Punnekkat, "Learning the parameters of periodic traffic based on network measurements," in *2015 IEEE International Workshop on Measurements & Networking (M&N)*, Oct 2015, pp. 1–6.

[20] J. Farkas, S. Haddock, and P. Saltsidis, "Software defined networking supported by ieee 802.1q," *CoRR*, vol. abs/1405.6953, 2014.

[21] J. Åkerberg, M. Gidlund, and M. Björkman, "Future research challenges in wireless sensor and actuator networks targeting industrial automation," in *2011 9th IEEE International Conference on Industrial Informatics*, July 2011, pp. 410–415.

[22] E. Lisova, E. Uhlemann, J. Åkerberg, and M. Björkman, "Towards secure wireless ttethernet for industrial process automation applications," in *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*, Sept 2014, pp. 1–4.

[23] E. Lisova, E. Uhlemann, W. Steiner, J. Åkerberg, and M. Björkman, "A survey of security frameworks suitable for distributed control systems," in *2015 International Conference on Computing and Network Communications (CoCoNet)*, Dec 2015, pp. 205–211.

[24] E. Lisova, M. Gutiérrez, W. Steiner, E. Uhlemann, J. Åkerberg, R. Dobrin, and M. Björkman, "Protecting clock synchronization - adversary detection through network monitoring," *Journal of Electrical and Computer Engineering*, April 2016.