Realtime Publish/Subscribe for the Industrial Internet of Things

Eike Schweissguth*, Helge Parzyjegla†, Peter Danielis†, Gero Mühl†, and Dirk Timmermann*
*Institute of Applied Microelectronics and Computer Engineering
†Institute of Computer Science
University of Rostock, 18051 Rostock, Germany
{eike.schweissguth, helge.parzyjegla, peter.danielis, gero.muehl, dirk.timmermann}@uni-rostock.de

***Abstract*—Publish/subscribe communication offers outstanding flexibility, but lacks the predictability required for guaranteeing (hard) realtime constraints in industrial environments such as future smart factories. In this paper, we review existing publish/subscribe systems, identify their shortcomings, and derive research objectives for a converged realtime network infrastructure for the Industrial Internet of Things (IIoT) that is based on publish/subscribe and Software-Defined Networking (SDN).**

I. Introduction

The seamless integration of applications, machines, processes, and people holds a disruptive potential for innovations that will both ease our daily lives and revolutionize industry, trade, and economics in future. Although precise details are up to mere speculations, the inherent fundamental characteristics can already be derived from current trends today. It is assumed that the working environment in a future *smart factory* is based on comprehensive digitalization, thorough interconnectedness, and maximum flexibility in terms of space, time, and organization [8]. This is required to provide the foundation for combining the empathy and creativity of humans with the power and precision of robots and the capability of *Artificial Intelligences* (AIs) to control and optimize complex production processes and whole value chains, respectively [4].

To facilitate such a tight cooperation of humans, machines, and AI, the network infrastructure of a smart factory has to support a wide range of applications and meet their individual requirements. These do not only include conventional *Quality of Service* (QoS) aspects such as bandwidth, latency, and jitter; but also properties like flexibility, reconfigurability, and scalability. Potential applications in a smart factory, for example, may range from control of production processes having hard real-time and fault-tolerance constraints, over augmented reality environments requiring large bandwidths and low latencies to best effort traffic which has to connect operational information systems and office information technology in a light-weight, easily reconfigurable manner.

Publish/subscribe communication has shown to flexibly support such versatile interaction patterns and, embodied in protocols such as MQTT, CoAP, and AMQP, it has become an integral part of the *Internet of Things* (IoT) [11]. But for its application in an industrial environment, publish/subscribe still lacks assured guarantees for hard realtime constraints. However, with the advent of *Software-Defined Networking* (SDN) and new realtime standards for networking such as *Time-Sensitive Networking* (TSN), this is about to change.

In this paper, we outline a converged network infrastructure for the *Industrial Internet of Things* (IIoT) in a smart factory. In particular, we aim to support publish/subscribe communication offering flexible $m{:}n$ multicast interactions at the network layer in which publishers and subscribers may send and receive with multiple data rates (*multi-rate*) and messages have realtime constraints of varying strictness (*multi-class*), e. g., hard deadlines, reserved bandwidth, and best effort. In the following, we first summarize the state of the art to pinpoint weaknesses and identify open issues. Based on our analysis, we outline research steps and questions that need to be approached and solved on the way towards a future IIoT infrastructure that is ready for production.

II. Summary of Related Work

Publish/subscribe has been proven successful for *enterprise application integration* (EAI) for many years [3] and is already widely used in the IoT [11]. Within these application domains, publish/subscribe middleware implementations evolved to meet diverse requirements such as offering persistent messages, durable subscriptions, and delivery guarantees, or coping with networks that have low bandwidths, long delays, and high error-rates. However, realtime constraints are usually not considered. With the *Data Distribution Service* (DDS) [5], a specification was designed for low-latency publish/subscribe communication that explicitly addressed periodic tasks with deadlines and latency budgets, too. The *Real-Time Publish/Subscribe* (RTPS) protocol ensures the interoperability of DDS implementations, but does not consider the lower network layers. Thus, it actually lacks the means to guarantee the deadlines and enforce the latency bounds. However, the latter can be achieved, even without special wire protocols, by reserving sufficient network resources. In [2], SDN is used to reserve bandwidth on links and buffer capacity on switches for each data flow. Applying the Network Calculus, bounds for the maximum latency can be derived and guaranteed. Nevertheless, guaranteeing very low isochronic cycle times is difficult without synchronization when compared to approaches based on scheduled and synchronized time-slots.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) as part of project RTiPS-4-IIoT (project no. 410407136).

On the contrary, Industrial Ethernet solutions usually support hard realtime constraints of meticulous strictness, but often lack flexibility and scalability. An exception is the *Flexible Time-Triggered* (FTT) paradigm [7] that allows to adapt the time-slot schedule from cycle to cycle and, thus, enables new (periodic as well as aperiodic) realtime data flows to be added at runtime. The *Hard Real-Time Ethernet Switching* (HaRTES) architecture [10] is implemented using an FPGA-based Ethernet switch with a custom firmware. The switch takes over the role of an FTT master and enforces an FTT schedule in the network. Although more recent work added a basic group communication mode [9], flexible $m{:}n$ publish/subscribe communication has not been achieved yet.

Bundled under the term *Time-Sensitive Networking* (TSN) [1], a set of Ethernet standards is currently developed targeting realtime communication in industrial environments. In particular, TSN standardizes the means to enforce bandwidth reservations and realtime schedules making Industrial Ethernet switches from different vendors interoperable. Recently, the *Open Platform Communications Unified Architecture* (OPC UA), designed as a service-oriented architecture for the description and exchange of IIoT machine data, has been supplemented with a modern publish/subscribe communication mode [6]. The combination with TSN to enforce QoS and realtime constraints looks very promising, although adequate routing and scheduling algorithms as well as established implementation approaches are still missing.

III. Research Objectives

With the standardization of TSN and the flexibility gained by SDN, we believe that the foundations for actually realizing hard realtime publish/subscribe communication are given. Nevertheless, major challenges still need to be solved in order to combine the predictability, flexibility, and scalability in an industrial communication environment. In particular, the following four research questions need to be addressed.

1. Modeling of Realtime Publish/Subscribe. Publish/subscribe interactions are often indirect, complex, and involve multiple participants each having individual requirements. Nevertheless, it is essential that communication interrelationships are concisely modeled (including a precise specification of all requirements) in order to be able to give deterministic guarantees. To achieve this, current publish/subscribe models must be improved to cover the relevant realtime aspects (e. g., publication and consumption rates, latency bounds, deadlines).

2. Planning, Routing, and Scheduling. For the realtime publish/subscribe traffic, communication schedules for participants and links have to be created and required network paths have to be reserved for occurring data flows. Assuming an appropriate and concise model, general (optimization) problem solvers (e. g., CSP solvers, ILP solvers) can be used for finding solutions for network routes and schedules. Nevertheless, specific and optimized algorithms and heuristics are still needed to reduce the computation time to an acceptable level in order to be used by a network controller at runtime.

3. Seamless Dynamic Reconfiguration. With participants and data flows being dynamically added and/or removed at runtime, it is necessary to reconfigure the network without interrupting existing flows and violating given realtime guarantees. To achieve this, potentially complex reconfigurations have to be broken down into a series of atomic reconfigurations that can be carried out in a seamless manner. Unfortunately, established reconfiguration strategies are still missing in both the publish/subscribe and the realtime domain.

4. Scaling. Finally, it is necessary to scale developed algorithms and strategies to be also applicable to larger networks that, for example, span multiple production facilities. In particular, this requires adequate abstractions and structuring means to decompose complex networks into smaller domains that can be better planned, scheduled, and reconfigured with the approaches above. Subsequently, partial plans and schedules need to be seamlessly merged into an overall solution.

IV. Conclusions

In this paper, we motivated a converged realtime communication infrastructure for the IIoT of a future smart factory based on publish/subscribe and SDN. The major challenge is to combine the flexibility, dynamics, and scalability of publish/subscribe with the predictability and guarantees required in realtime environments. We analyzed shortcomings of existing publish/subscribe approaches and identified four primary research objectives for future work: (i) adequate communication models, (ii) novel algorithms for planning and scheduling, (iii) strategies for seamless reconfigurations at runtime, and (iv) abstractions and means for scalability.

References

[1] J. Farkas, L. Lo Bello, and C. Gunther. Time-sensitive networking standards. *IEEE Communications Standards Magazine*, 2(2):20–21, June 2018.

[2] J. W. Guck and W. Kellerer. Achieving end-to-end real-time quality of service with software defined networking. In *3rd IEEE International Conference on Cloud Networking (CloudNet 2014)*, pages 70–76. IEEE, Oct. 2014.

[3] G. Hohpe and B. Woolf. *Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions*. Addison-Wesley, Boston, MA, USA, 2003.

[4] J. Köbler. Kollege 4.0. *Audi Dialoge Smart Factory*, pages 40–43, 2017.

[5] Object Management Group, Inc. Data distribution service for real-time systems (DDS), version 1.2. Needham, MA, USA, Jan. 2007.

[6] OPC Foundation. OPC unified architecture part 14: PubSub. Scottsdale, AZ, USA, Feb. 2018.

[7] P. Pedreiras and L. Almeida. The flexible time-triggered (FTT) paradigm: an approach to QoS management in distributed real-time systems. In *17th IEEE International Parallel and Distributed Processing Symposium (IPDPS 2003)*. IEEE, Apr. 2003.

[8] H. Reil. Smart faction. *Audi Dialoge Smart Factory*, pages 26–31, 2015.

[9] G. Rodriguez-Navas and J. Proenza. A proposal for flexible, real-time and consistent multicast in FTT/HaRTES switched ethernet. In *18th IEEE Conference on Emerging Technologies and Factory Automation (ETFA 2013)*, pages 1–4. IEEE, Sept. 2013.

[10] R. Santos, R. Marau, A. Vieira, P. Pedreiras, A. Oliveira, and L. Almeida. A synthesizable ethernet switch with enhanced real-time features. In *35th Annual Conference on IEEE Industrial Electronics (IECON 2009)*, pages 2817–2824. IEEE, Nov. 2009.

[11] M. B. Yassein, M. Q. Shatnawi, and D. Al-zoubi. Application layer protocols for the internet of things: A survey. In *2016 International Conference on Engineering & MIS*. IEEE, Sept. 2016.