

Position Paper on Time and Event-triggered Communication Services in the Context of e-Manufacturing

Eduardo Tovar, Luis Miguel Pinho
IPP-HURRAY Research Group
Polytechnic Institute of Porto, Portugal
{emt, lpinho}@dei.isep.ipp.pt

Luís Almeida
DET / IEETA
Universidade de Aveiro, Portugal
lda@det.ua.pt

Abstract

Modern factories are complex systems where advances in networking and information technologies are opening new ways towards higher efficiency. Such move is being driven by market rules with ever-increasing competition levels, in search for faster time-to-market, improved process yield, non-stop operations, flexible manufacturing and tighter supply-chain coupling. All these aims present a common requirement, i.e. a real-time flow of information, from the plant-floor up to the management, maintenance, suppliers and clients, to support accurate monitoring and control of the factory. This stresses the importance achieved by the communication infrastructure in modern manufacturing industry.

This paper presents the authors view concerning the current trends in modern factory communication systems. It addresses the problems of seamlessly integrating different information flows with diverse requirements, mainly in terms of timeliness. In this aspect, the debate between event-triggered and time-triggered communication is revisited as well as the joint support for both types of traffic. Finally, a view of where factory communication systems are moving to is also presented, showing the impact of open and widely available technologies.

1 Motivation

The most strategic advantage of any manufacturing enterprise today is information. Whether the challenge is

This work was partially supported by the Portuguese Government through FCT (projects CIDER POSI/1999/CHS/33139 and MethoDES POSI/2001/37334) and by the European Commission (accompanying measure ARTIST, IST-2001-34820).

faster time-to-market, improved process yield, non-stop operations or tighter supply-chain coupling, information is the key. The plant is crucial to every manufacturing enterprise. It is the place where value is created. And technology is not new to the plant floor. Highly sophisticated combinations of relays and switches, programmable controllers, drives and sensors have long operated the lines, conveyors, and machinery required to manufacture the goods.

As industries consolidate and restructure, some companies are choosing to remain as the “producers” of goods while others are positioning themselves downstream in the supply chain to be the “marketers” of those same goods. In almost every industry, tightly integrated supply chain models are emerging. In these new models, connection of the plant floor to the broader supply chain is essential, and information access is more critical than ever. The Internet and e-Commerce have simply accelerated this trend towards the e-Manufacturing [1].

The factory-floor is the starting point for greater information connectivity. Computer-based factory-floor controls for manufacturing machinery, materials handling systems and related equipment generate a wealth of information about productivity, product design, quality and delivery. Factory-floor networking arises as the key to unleashing this information in a cost-effective manner.

Over the last decade, the evolution of the manufacturing industry has assimilated “contemporary” information technologies, such as the PC and open standards for the factory-floor networking. In fact, the application of information technologies has evolved from relatively passive data collection and reporting roles to feedback control and diagnostics applications. This context unveils the role played by factory-floor networking in modern industrial automation systems.

It is interesting, however, to notice that the use of communication networks at the factory-floor is more recent than at the office environment. One of the reasons for this delay was that manufacturing systems usually depend on being able to sample input data at equally spaced points in time [2] and this feature was not easily fulfilled using early office-room networks. This led to the fieldbus concept, the first step on the road to networked industrial automation systems. Fieldbus networks aim at the interconnection of sensors, actuators and controllers, being generally adapted to convey periodic and aperiodic, real-time and non-real time traffic. Different technologies are available in the market, being PROFIBUS [3], DeviceNet [4], WorldFIP [5], and TTP [6] just a few but representative examples. This diversity is partially justified by the diversity in the requirements imposed by the different targeted applications.

In the era of the Internet, however, factory-floor communication systems must also better explore commercial information technologies. This includes COTS operating systems, TCP/UDP/IP based applications (XML, Java, etc.) [7-8] and general-purpose communication networks such as Ethernet [1,9]. The reason is not only to integrate factory-floor operations to work seamlessly throughout a typical plant environment, but to also make that control information transparent throughout the overall enterprise. These “open” technologies present a strategic advantage to address the control needs of the plant and the connectivity to achieve that transparency.

In this trend, networks must deliver an efficient means to exchange data for precise control, while supporting non-critical systems and device configuration at start-up and monitoring during run-time. Networks must also provide the critical link for collecting data at regular intervals for analysis and feedback control. An integrated architecture benefits from a common set of advanced network services and interfaces optimized for timeliness control, configuration and collection of data, with seamless communications up and down the architecture, allowing access to any part of a system from any location. Ethernet/IP (IP stands for Industrial Protocol) [10-11] is one example of how factory-floor networks are migrating to general-purpose information technologies, thus providing a flexible exchange of the critical production data for the benefit of the greater enterprise.

Advances in networking and information technologies are transforming factory-floor communication systems into a mainstream activity within industrial automation. It is now recognized that future industrial computer systems will be intimately tied to real-time computing and to communication technologies. For this vision to succeed, complex heterogeneous factory-floor communication

networks (including mobile/wireless components [12-13]) need to function in a predictable, flawless, efficient and interoperable way. The following sections revisit the debate between event-triggered and time-triggered communication as paradigms for supporting the desired integration of different networks and types of traffic. Finally, the use of widely available networking technologies, namely Ethernet, in factory communication systems is discussed, highlighting its advantages and current problems.

2 Time vs. Event-Triggered Services

The integration of the computing technology with the underlying communication infrastructure is turning out to be a key feature of large-scale distributed computer-controlled systems.

However, the role of the communication infrastructure should not be exacerbated to the point where the model of the network restricts the design of the system, leading to low design flexibility and potentially low efficiency.

Typically, communication networks for distributed computer-controlled systems must deliver both synchronous (take it as time-triggered – TT) as well as asynchronous (take it as event-triggered – ET) communication services under timing constraints. Some authors [14] claim, and to some extent correctly, that TT vs. ET is not a fundamental question, since there is not enough evidence of fundamental issues separating them.

However, the fact is that one of the main aspects of the network that is often discussed, and seen as a constraint to the design of control applications, concerns actually the TT vs. ET debate. This inexhaustible discussion leads to the impression that a system is either TT or ET, and therefore should be either supported by TT or ET architectures. It is just an impression.

This discussion gained extra relevance with the emergence of TTP [6] in the early 90s. TTP highlighted the advantages of the TT paradigm in real-time communication systems. The impact was so relevant that the ISO Technical Committee TC22/SC3/WG1 set up in 1999 a task force (TF6) to work out a definition of a CAN-based time-triggered standard, TT-CAN [15]. On the other hand, despite not defining application services CAN allows direct send and receive primitives to be easily implemented over the respective MAC in an ET fashion. This means that nodes access the bus in a completely asynchronous way, using another mechanism to resolve eventual collisions. For this reason, CAN is normally referred to as following the ET paradigm. A similar observation can be made concerning Ethernet, too.

One of the considered advantages of event-triggered communication is that it is more efficient in dealing with the communication resources. However, when worst-case situations are considered, that efficiency is not verified. Since events are asynchronous by nature, a typical worst-case assumption is that all events that must be handled by the system will arrive simultaneously. To cope with such situation in a timely fashion, the required amount of resources (e.g. network bandwidth) is very high.

On the contrary, the time-triggered approach forces the communication activity to occur at pre-defined instants in time at a rate determined by the dynamics of the environment under control. One of the features of this approach is that it allows relative phase control among the streams of messages to be transmitted over the communication system. By using this feature, messages of different streams can be set out-of-phase allowing a reduction on the number of messages that become ready for transmission simultaneously.

This feature is responsible for one of the most important properties of time-triggered communication; that is, the support for composability with respect to the temporal behavior [16]. This property assures that, when two subsystems are integrated to form a new system, the temporal behavior of each of them will not be affected.

This does not hold true for event-triggered communication. In this case, the level of contention at the network access that each subsystem feels before integration is always increased upon integration due to the traffic generated by the other subsystem.

Furthermore, the relative phase control allowed by the time-triggered approach may lead to two other positive effects. Firstly, it improves the control over the transmission jitter felt by periodic message streams. Secondly, it supports higher network utilization with timeliness guarantees. Therefore, when considering worst-case situations the time-triggered approach is more resource efficient than the event-triggered one. However, when considering average-case situations, time-triggered communication is considerably greedy when compared to event-triggered one.

Consequently, by dimensioning a system according to its worst-case requirements, as typical in hard real-time systems, the time-triggered approach tends to be less expensive. Nevertheless, since the average network utilization of event-triggered systems is normally lower, such systems can easily support other sorts of communication with less stringent or no timing constraints without any additional cost. This fact can have a positive impact on the overall efficiency of the communication system utilization, thus leading to reduce its exploitation costs.

Apart from the above considerations on network

utilization, it is commonly accepted [17-18] that time-triggered communication is well adapted to control applications that typically require regular transmission of state data with low, or bounded, jitter (e.g. motion, engine, temperature and position control). On the other hand, event-triggered communication is well adapted to the monitoring of alarm conditions that should occur sporadically and seldom, and also to support asynchronous non-real-time traffic, e.g. for system management.

Finally, there are other aspects typically referred in this context such as the capacity of the TT model to support fast detection of omissions, and consequently of silent failures, and its suitability to support formal analysis by means of synchronous languages. However, concerning these two particular aspects, ET systems can also detect remote failures, typically using timeouts, and, in general, can also be formally analyzed, e.g. with respect to the network-induced delay.

3 Combined TT and ET Support

Many applications do require joint support for all these sorts of traffic and thus, a combination of both paradigms is desirable to benefit from the best of the each world. However, such combination must enforce temporal isolation between both types of traffic so that the transmission instants of the TT traffic are not disturbed by the asynchrony of the ET traffic.

There are two alternative approaches for integrating both types of communication services, with temporal isolation, in a communication system. In one approach, bandwidth is allocated to each of the two types of traffic. The bus time becomes an alternate sequence of time-triggered and event-triggered phases. On the other approach, event-triggered messages use pre-allocated slots of a time-triggered communication protocol. Typically this requires the use of dual-stack communication architecture, on top of a single medium access protocol.

There are already in the market communication networks that implement the first approach. A classical is WorldFIP [5]. However, since this fieldbus uses a MAC protocol based on centralized arbitration, the handling of event-triggered (aperiodic) traffic is rather inefficient requiring a considerable amount of bandwidth to allow the master node (arbitrator) to become aware of aperiodic requests. In the Foundation Fieldbus [19], a somewhat similar scheme is used. FlexRay [20], a new protocol that is being developed by a consortium of carmakers, also enforces temporal isolation between the two types of traffic. In this case, the access scheme for the aperiodic traffic is based on wait times (mini-slotting), a technique that is more bandwidth efficient than centralized

arbitration but may impose considerable access delays to messages with low priority identification numbers.

In many other fieldbus systems, that do not use different bus-time phases, it is still possible to specify cyclic time-triggered data exchanges but with no temporal isolation from the event-triggered traffic. One example is PROFIBUS. For this communication protocol, a dispatcher sub-layer [21] was proposed to add temporal isolation assuming the setting of network parameters in such a way that the token is never late [3].

The time-triggered version of CAN (TT-CAN) [15] gives such a solution, by turning the bus into a sequence of different windows, through a pre-defined time schedule. In this solution, Exclusive Windows are reserved for particular messages, while Arbitration Windows use the standard non-destructive arbitration method of event-triggered CAN.

The Flexible Time-Triggered protocol for CAN (FTT-CAN) [22] is an example of a solution that provides joint support of event- and time-triggered communication. As in TT-CAN, FTT-CAN supports alternate phases of event-triggered and time-triggered communication within a bus cycle (called Elementary Cycle). However, contrarily to TT-CAN, the protocol supports dynamic communication requirements, by using a centralized scheduling with on-line admission control, allowing the network schedule to evolve over time. The FTT approach was also extended to Ethernet networks [23], providing them with flexible and efficient support for hard real-time traffic.

Providing an event channel on top of a time-triggered protocol, by reserving slots for event-triggered messages, allows not changing already existent protocols, maintaining their composability and isolation properties. Examples of this approach can be found, for instance, in the CAN and TCP/IP emulation on top of TTP [24,25]. However, this solution can be very inefficient since the sporadic nature of events does not map to the periodic nature of time-triggered slots, and unused slots are not made available for other traffic [26].

4 What's on About Ethernet Technologies?

Nowadays, computer-controlled systems are generally complex and can be found in areas such as factory automation, process control, robotics, automotive systems, avionics and space applications.

Furthermore, future embedded systems are expected to be more dynamic in nature than the current static systems. These systems will have varying time and resource demands and must deliver dependable and adaptable services with guaranteed temporal qualities.

Changes in the environment may cause events to arrive more frequently or even trigger new computational activities as a reaction, hence affecting the activation rates and the load distribution in the system. A change in the system structure may happen due to maintenance, network re-configuration, new hardware installation, or component failures. In both cases, the system should ensure a minimum level of performance for the most critical services, while providing the best possible quality of service for the non-critical ones [9].

Although Ethernet has been used primarily as an information network, there is a strong belief that some very recent technological advances will enable its use in dependable applications with real-time requirements. First of all, Ethernet has evolved to operate at 100Mbps and 1Gbps (with 10Gbps in development). Therefore, it provides a high bandwidth, when compared to that of CAN (1Mbps) or of PROFIBUS (12Mbps), just to give a few examples. Also, the mechanisms associated with switched-Ethernet (such as priorities, flow control, spanning tree, port trunking, virtual LANs, etc.) seem to be promising to enable the support to flexible distributed applications [9], whilst guaranteeing the determinism and reliability requirements of hard real-time applications. Moreover, the recent standardization of wireless Ethernet [26] introduces the support to time-bounded delivery services, allowing the integration of wireless links in the communication infrastructure.

These advances enable to consider Ethernet as a complex network assembly, built as a set of interacting components (bus/star topologies and wired/wireless links), all sharing the same higher-level interfaces. Thus, it is a potential candidate for a communication infrastructure of advanced distributed computer-controlled systems.

However, by themselves these advances do not allow to achieve a deterministic and reliable communication infrastructure. Ethernet is an event-triggered network, with very few mechanisms for guaranteeing the temporal and isolation requirements of distributed computer-controlled systems. The use of these mechanisms must be assessed, and additional mechanisms must be integrated in order to provide the support for both time-triggered and event-triggered communication, consistent message delivery, isolation and network composability, also providing flexibility in order to allow dynamic accommodation of soft or non real-time traffic. The approach for the development of simultaneous (and integrated) transmission of time- and event-triggered messages over Ethernet is still an on-going effort, and addresses several issues (to mention a few): the need for a global time frame; the definition of an integration model;

efficient bandwidth allocation; provision for dynamic flexibility with online admission control; provision for temporal isolation and composability in heterogeneous networks and the support to consistent message delivery and component redundancy.

One of the most important issues is the provision of a global time base, which can be used as a reference for timeliness communication. Although common Ethernet networks do not provide clock granularity suitable for distributed computer-controlled systems, approaches for time synchronization over switched-Ethernet networks exist [27], and may be used.

The provision of an integrated ET/TT traffic approach in Ethernet has already seen some relevant steps [23,28]. The use of the switch features can improve the time-triggered support, since there is no need for tight control on the message transmission instants [28]. In the situation of controlled loads, switches prevent the destructive collision of the shared Ethernet approach, thus allowing a more efficient management of the available bandwidth.

On the other hand, the use of time-triggered communication together with on-line admission control for event-triggered soft or non real-time traffics can improve the limitations currently seen in switches, since traffic control is also performed within the system nodes. Also, priority scheduling can be performed beyond the priority levels currently available in IEEE802.1D.

Concerning the use of wireless links, several impairments are pointed to the wireless version of Ethernet. The standard uses a large bandwidth in comparison to the data rates delivered, frequency planning is required and there is no QoS support. Therefore, there is a great deal of work to be done if IEEE802.11 is to be used for real-time transmission. Even if the wireless links are only used to transmit soft or non real-time traffic, the integration of the wireless links with the wired infrastructure (which will support hard real-time) must be well defined, in order to guarantee the composability of the system.

5 Conclusion

Computer-based factory-floor controls for manufacturing machinery, materials handling systems and related equipment generate a wealth of information about productivity, product design, quality and delivery. Factory-floor networking arises as the key to unleashing this information in a cost-effective manner.

A manufacturing plant is today just a component of a tight supply chain enterprise, where the connection of the plant floor to the broader supply chain is essential, and information access is more critical than ever.

In this eManufacturing era, factory-floor communication systems must also better explore commercial information technologies. The reason is not only to integrate factory-floor operations to work seamlessly throughout a typical plant environment, but to also make that control information transparent throughout the overall enterprise. These "open" technologies present a strategic advantage to address the control needs of the plant and the connectivity to achieve that transparency.

We also believe that new user-computer interaction techniques including multimedia and augmented reality combined with, now more affordable, technologies like wearable computers, pen aware computing, voice interaction and wireless networks can change the way the factory personal works together with the machines and the information system on the factory-floor.

In fact, in the past few years the so-called "gadgets" like cellular phones, personal data assistants and digital cameras have become widespread even with less technological-aware users. However, the factory-floor itself seems to be hermetic to these changes... Probably because of fear that they could disturb the operation of time-critical distributed computer-controlled applications.

This gives the context to revisit the debate between event and time-triggered traffic support in communication networks. Both paradigms have their advantages and disadvantages and the way to go seems to be their seamless integration, profiting from both in a factory-floor communication infrastructure. As for the network technologies, recent advances in Ethernet, its high speed and its dominant position in the LAN market, make one wonder about the role of Ethernet technologies in future real-time distributed factory-floor applications. For those looking for a single technology to support factory-wide communication, Ethernet seems to be a serious candidate.

References

- [1] Rockwell Automation, "Making Sense of e-Manufacturing: a Roadmap for Manufacturers", Rockwell Automation White Paper, 2000.
- [2] Skeie, T., Johannessen, S and O. Holmeide, "The Road to and End-to-End Deterministic Ethernet", Proceedings of the 4th IEEE International on Factory Communication Systems (WFCS'2002), pp. 3-9, August 2002.
- [3] Tovar, E. and F. Vasques, "Real-Time Fieldbus Communications Using Profibus Networks", IEEE Transactions on Industrial Electronics, Vol. 46, No. 6, pp. 1241-1251, December 1999.
- [4] Rockwell Automation, "DeviceNet Product Overview", Publication DN-2.5, Rockwell Automation, 1997.
- [5] Almeida, L., Tovar, E., Fonseca, J. and F. Vasques, "Schedulability Analysis of Real-Time Traffic in

- WorldFIP Networks: an Integrated Approach", IEEE Transactions on Industrial Electronics, Vol. 49, No. 5, pp. 1165-1174, October 2002.
- [6] Kopetz, H. and G. Grünsteidl, "TTP - A Protocol for Fault-Tolerant Real-Time Systems", IEEE Computer, 27(1), January 1994.
- [7] Pereira, N., Pacheco, F., Pinho, L. M., Prayati, A., Nikoloutsos, E., Kalogeras, A., Hintze, E., Adamczyk, H. L. Rauchhaupt, "Integration of TCP/IP and PROFIBUS Protocols", WIP Proceedings of the 4th IEEE International Workshop on Factory Communication Systems (WFCS2002), August 2002.
- [8] Ferreira, L and E. Tovar, "QoS of IP Services in a Fieldbus Network: on the Limitations and Possible Improvements", WIP Proceedings of the 14th Euromicro Conference on Real-Time Systems (ECRTS02), June 2002.
- [9] Alves, M, Tovar, E., Fohler, G. and G. Buttazzo, "CIDER: Envisaging a COTS Communication Infrastructure for Evolutionary Dependable Real-Time Systems", WIP Proceedings of the 12th IEEE Euromicro Conference on Real-Time Systems (ECRTS2002), pp. 19-22, June 2000.
- [10] B. Batista, "Industrial Ethernet: Building Blocks for a Holistic Approach", WIP Proceedings of the 4th IEEE International Workshop on Factory Communication Systems (WFCS2002), August 2002.
- [11] A. Moldovansky, "Performance Optimization Recommendations for EtherNet/IP Networks", WIP Proceedings of the 4th IEEE International Workshop on Factory Communication Systems (WFCS2002), August 2002.
- [12] Alves, M., Tovar, E., Vasques, F., Hammer, G. and K. Roether, "Real-Time Communications over Hybrid Wired/Wireless PROFIBUS-based Networks", Proceedings of the 14th Euromicro Conference on Real-Time Systems (ECRTS'02), pp. 142-150, June 2002.
- [13] Dzung, D., Apneseth, C., Scheible, G, and W. Zimmermann, "Wireless Sensor Communication and Powering System for Real-Time Industrial Applications", WIP Proceedings of the 4th IEEE International Workshop on Factory Communication Systems (WFCS2002), August 2002.
- [14] P. Veríssimo, "Fundamental Questions in the ET vs. TT Debate? Please Look Elsewhere", Booklet of the NextTTA Workshop on the Integration of Event-Triggered and Time-Triggered Services, October 2002.
- [15] ISO/WD11898-4, "Road Vehicles – Controller Area Network (CAN) – Part 4: Time-Triggered Communication", December 2000.
- [16] H. Kopetz, "Real-Time Systems Design Principles for Distributed Embedded Applications", Kluwer Academic Publishers, 1997
- [17] Thomesse, J.-P. and M. Leon Chavez, "Main Paradigms as a Basis for Current Fieldbus Concepts", Proceedings of the International Conference on Fieldbus Technology (FET'99), pp. 2-15, September 1999.
- [18] Peraldi, M.A. and J.D. Decotignie, "Combining Real-Time Features of Local Area Networks FIP and CAN", Proceedings of the 2nd Int. CAN Conference (ICC'95), CiA – CAN in Automation, 1995.
- [19] IEC International Standard 61158, "Fieldbus standard for use in industrial control systems – Type 1: Existing IEC TS61158 parts3-6 (Foundation Fieldbus H1)", International Electrotechnical Committee, 2000.
- [20] F. Bogenberger, B. Müller, and T. Führer, "Protocol Overview," *Proc. FlexRay Int'l Workshop*, 2002; <http://www.flexray.com/htm/infws402.htm> (current July 2002).
- [21] Ferreira, L, Tovar, E. and S. Machado, "Scheduling IP TRaffic in Multimedia Enabled PROFIBUS Networks", Proceedings of the 8th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'2001), pp. 169-176, October 2001.
- [22] Pedreiras, P., Almeida, L. and J. Fonseca, "The FTT-CAN Protocol: Why and How", IEEE Transactions on Industrial Electronics, Volume 49, Number 6, December 2002.
- [23] Pedreiras, P., Almeida, L. and P. Gai, "The FTT-Ethernet Protocol: Merging Flexibility, Timeliness and Efficiency", Proceedings of the 14th Euromicro Conference on Real-Time Systems (ECRTS'02), pp. 153-160, June 2002.
- [23] R. Obermaisser, "Can on Top of TTP/C", Booklet of the NextTTA Workshop on the Integration of Event-Triggered and Time-Triggered Services, October 2002.
- [24] I. Majzik, "TCP/IP on Top of TTP/C", Booklet of the NextTTA Workshop on the Integration of Event-Triggered and Time-Triggered Services, October 2002.
- [25] Suri, N. and V. Claesson, "Approaches, Scenarios & Issues for Integrating ET/TT", Booklet of the NextTTA Workshop on the Integration of Event-Triggered and Time-Triggered Services, October 2002.
- [26] IEEE Std 802.11, "IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks", 1999 Edition (ISO/IEC 8802-11: 1999).
- [27] Johannessen, S., Skeie, T. and O. Holmeide, "Highly Accurate Time Synchronisation over Switched Ethernet", Proceedings of the 8th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'2001), pp. 195-204, October 2001.
- [28] Pedreiras, P. and L. Almeida, "Flexibility, Timeliness and Efficiency in Ethernet", Proceedings of the 1st Intl. Workshop on Real-Time LANs in the Internet Age (RTLIA'02), June 2002.