The role of CAN in the age of Ethernet and IOT

Christian Schlegel, HMS Industrial Networks

CAN technology was developed in the 1980s and became available in 1987, just as other industrial fieldbus systems like PROFIBUS or INTERBUS entered the stage of industrial communication. Beside the fact that CAN is a success in the automotive industry and used in all types of cars today, it has also made its way in many other industrial areas. About 15 years ago, new technologies based on Ethernet started to emerge, with appealing and sometimes outstanding features. Some six years ago Ethernet also started to find its way into automobiles. Today, other new communication technologies are showing up on the horizon driven by the omnipresent Industrial Internet of Things. But even now, 30 years after their introduction, these "classic" fieldbus technologies are still alive – with varying success. Since CAN was initially developed with a focus for use in automobiles, CAN has certain features that still make it the best choice for many applications in automobiles and industrial areas – even when compared to the newer technologies.

This paper discusses why CAN is still a valid or even better choice for certain application areas than Ethernet-based technologies, not just focusing on the advanced features provided by the enhanced capabilities of CAN FD but also highlighting how these applications benefit from the features of "classic" CAN.

Looking back into history … … when CAN was born

In the early 80s, the increased use of electronics in all areas, especially in automation, required new communication technologies with more bandwidth and features. As a consequence some of the larger automation equipment manufacturers started to develop new technologies. These activities resulted in the publication of several new fieldbus technologies like PROFIBUS, INTERBUS or Sercos in the second half of the 80s. All of them were still based on a serial communication principle but with considerably increased wire speed, more features and – due to the typical architecture of an automation system — all of them were based on a master/slave communication principle. These started to replace the existing serial communication technologies.

Similar to the automation world, more and more electronics were introduced in passenger cars in form of Electronic Control Units (ECUs). This also imposed the requirement to transmit data between these ECUs but also to connect sensors and actuators to them. Using conventional cabling concepts would have led to a huge increase of cables and consequently of the weight of a car. There was also a need for a new way of exchanging data in a reliable and secure way between these ECUs, sensors and actuators.

As there are many diverse function areas inside a car like motor management, breaks, chassis control or passenger compartment and certain functions are optional depending on the car series or the customers' wallet, there is no clearly structured communication architecture inside a car like it is in the automation world. At a first glance, the communication structure of a car seems to be chaotic. In addition, the electric world inside a car is not as stable as in the automation world. For example there are many interferences from the engine which disturb the communication. Moreover, shielding of cables has to be avoided in order to safeguard weight. Consequently, communication strategies and technologies from the automation world were not applicable in cars and a new approach was necessary.

In the early 80s, a group of people at Bosch were the first to investigate the existing serial communication systems regarding possible use in passenger cars. However, none of these systems were able to fulfill the requirements and the development of a new serial communication system was started in 1983. Finally, in 1986, at the SAE congress “Controller Area Network” was introduced to the public. Only one year later, the first CAN controller chip, the 82526 was presented by Intel and shortly after Philips Semiconductor presented the 82C200.

Ingenious features made CAN different

Designed for use in passenger cars, the inventors gave CAN several clever and sophisticated features to which none of the new upcoming industrial fieldbus systems could compete with.

In order to understand why CAN is sometimes a better choice compared to Ethernet, some of the most important features need to be understood:

- The first and most outstanding feature to mention is the bus access method. CAN uses the Carrier Sense Multiple Access (CSMA) principle which was not a new principle: for example, it was already used by Ethernet. With this principle, all nodes of a network which want to transmit a message, listen to the network and if there is no transmission, they can start to transmit. This leads to collisions in case more than one node start to transmit. Consequently the messages are destroyed and all transmitting nodes have to resolve this situation by abandoning their transmission for a random time which might lead to longer delays, especially when there are many nodes which want to transmit data. In order to avoid this situation, the arbitration mechanism of CAN avoids collisions and allows one node of all the nodes which start to transmit their message at the same time to continue its transmission and all others immediately start to receive the message. As all CAN messages have an inherent message priority, the most prioritized message always wins the arbitration.
- In combination with the error signaling mechanism, CAN ensures that the data transmitted in the network is always consistent in all nodes connected to the network. Each node which identifies an error in a message immediately transmits a special error signature (error frame) destroying the currently transmitted message and making sure that all other nodes in the network also consider the currently transmitted message to be erroneous and, as a consequence, reject it. Error situations are e.g. checksum error, message frame delimiter errors (special bits of a message always have to have a defined value), bit stuffing error (after 5 consecutive bits with the same value, an additional inverted bit is inserted into the message). An error can also be signaled by the transmitting node if the currently transmitted bit seen on the network is different from the bit information the node actually transmits. If there is an error in a message transmitted on the bus, the transmitting node immediately stops the transmission and restarts a new transmission but with re-executing the arbitration procedure. If there is a message with higher priority at another node which shall be transmitted, the message with the higher priority will win the arbitration and be transmitted first.
- The arbitration mechanism including message prioritization and error signaling is only possible because of the physical representation of a bit on the bus line. A bit is transmitted in form of a differential signal on two twisted lines, ISO11898-2 (High Speed) defines a differential signal of 0V as recessive bit and a differential signal of 2V as a dominant bit. The symmetric transmission makes CAN immune to common mode interferences and the twisted pair lines compensate electromagnetic interferences. This makes CAN very resistant against external disturbances in general. For special purposes, fiber optic cables can also be used.

CAN and its way into industrial applications

The first company to use CAN in a mass produced car was Mercedes who started to use CAN in the S class in 1991.

At the same time as CAN was introduced, also the “classic” fieldbus systems like PROFIBUS, Sercos or INTERBUS were being introduced, especially for industrial applications. Although the major companies driving CAN were from the automotive market and hence have been focused on passenger cars, there has been a few other, rather small companies which started actively promoting CAN for industrial applications. Some of these pioneers were. Kvaser AB, I+ME GmbH, Softing and STZP (Steinbeis Transfer Center Process Automation, which became IXXAT Automation GmbH in 1998 and was acquired by HMS Industrial Networks AB in 2013). The distinctive features of CAN were also very interesting for industrial applications and so very early first applications were developed using CAN as a dedicated communication network. These applications used an own way of defining which data should be used and how data was to be transmitted within the CAN messages. Some of these first applications were developed by STZP for customers paving the way for CAN. These included a coil winding machine, a machine for validating and counting paper money, a security door control system for banks, a CAN-controlled endoscope system or a cow milking system with decentralized controllers for milking, feeding and weighting (see fig. 1)

Figure 1: CAN-controlled cow milking and feeding system developed 1992 (Source: STZP / HMS)

Looking at these application examples and others from industrial CAN pioneers, CAN was used very early in different application areas and markets — and not only focusing on industrial factory automation like the “classic” fieldbuses. For sure, this was the result of ease of use, reliability and the simple concept of CAN.

Because CAN was a really new technology, there was also a need to connect PCs (at that time running MS-DOS) to a CAN network for either running a control application or for running an analyzer software which allowed users to monitor what was happening on the CAN bus.

This lead to the development of first CAN products with focus on industrial applications. Like CAN interface boards as shown in figure 2 or the very first CAN repeater as shown in figure 3.

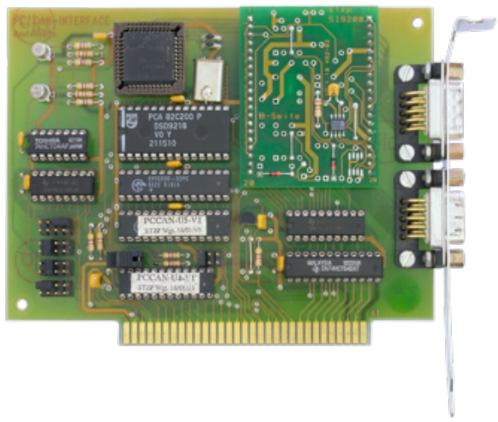

Figure 2: First PC CAN interface board with ISA bus hosting the Philips 82C200 (BasicCAN) and already the Intel 82527 (FullCAN) CAN controller chips (Source: STZP / HMS)

Figure 3: First CAN Repeater presented at the INTERKAMA fair in Dusseldorf in 1992 (Source: STZP / HMS)

In order to broaden the know-how, acceptance and use of CAN, a supporting ecosystem is necessary, consisting of product suppliers, consultants, system integrators and the machine builders implementing it. As learned in the first Keynote paper from Holger Zeltwanger, this requires much more than only a system and technology which allows to transmit data. An application layer with device and application profiles based on commonly agreed definitions and specifications becomes indispensable. The work to do that started immediately when CAN in Automation was founded in 1992 and some years later, the first higher layer standard for CAN became available. Today CAN in Automation has the most comprehensive set of standards for applications and device profiles compared to all other industrial communication technologies.

On the other hand, the passenger car manufacturers have been applying a completely different approach for the use of CAN in cars. Each manufacturer started to define an own catalog specifying for each data (or signal as they call it) which ECU transmits the data, which ECUs receive the data, and how often the data shall be transmitted. Years later, additional common definitions started to became standards e.g. in AUTOSAR for diagnostic purposes.

Ethernet - A threat for CAN?

Around the year 2000, Ethernet entered the stage in the industrial automation world. Already some years before, there was a discussion about using Ethernet in industrial automation systems. As a consequence, the first organization targeting common specifications for applying Fast Ethernet in industrial automation systems was founded in 1999 — IAONA. In 2001, Rockwell Automation launched EtherNet/IP as a first real industrial Ethernet specification and then other standards were introduced one after the other: Powerlink V1 in 2001, EtherCAT and PROFINET in 2003, Sercos III in 2004 and POWERLINK V2 (today's POWELRINK) in 2006. This was only possible because all the companies behind these standards started working on these standards already by the end of the 90ies.

Why was everyone looking into Ethernet as replacement of the fieldbus networks? There are some aspects which are obvious: Ethernet is used widely in office networks, public infrastructure and the Internet. Consequently Ethernet components like PHYs or Transceivers are produced in huge quantities which brings the price down. Furthermore, Ethernet brings the possibility of transmitting messages with a larger amount of data in one message and in general a much larger bandwidth due to the fact that it runs at 100 Mbit/s whereas CAN and the classic fieldbuses are limited in speed and data per message. Furthermore, the number of nodes in an Ethernet network is almost unlimited. This is very interesting especially in the area of factory automation: here we have large machines requiring a considerable amount of data to be transmitted for controlling the machine. The machines can be large and so the network extension becomes large. Also networks between machines on the factory floor interconnecting machines become more and more important and also require a high data capacity and long network extensions. Finally, as machines become more and more sophisticated, they impose high demands on motion control. This results in requirements to much faster cycle times and also more data to be transmitted in shorter times guaranteeing real-time behavior.

This may apply for the use of Industrial Ethernet in larger machines, fast running machines and basically in the area of factory automation. But what about applications and markets which do not have such requirements? Like small machines (ticket machines, vending machines, labelling machines,…), machines in medical, laboratory automation, test & measurement, mobile machines or utility and farming machines (figure 4).

Figure 4: Today's application areas of CAN based communication networks

Would it make sense to deploy Ethernet-based communication networks in general also in these areas?

The answer is: Definitely not, and there are good reasons for staying with CAN as long as the available data capacity, bandwidth and real-time capability of CAN meet the requirements of these applications:

- Ethernet and the derived Industrial Ethernet systems are based on a point-to-point connection. This means that active components in form of switches are required to interconnect nodes. Switches can be implemented with two external ports on a device allowing daisy-chain (line) topologies or as external switches allowing star topologies. If one device with an integrated switch or if an external switch device fails, larger parts of the network may not be accessible anymore. CAN on the other hand is using a passive connection method. A CAN node can be connected anywhere to the cable since CAN allows line topologies and star topologies with short branches. In case of a daisy chain of CAN nodes, the CAN cable is routed in a passive way through the CAN node (just directly connecting the cables inside the device). Failing of such a device will not result in losing connection between all the other nodes in the network.
- In automation and control applications, the data to be transmitted typically consists of short data packages (e.g. command, parameter, status) with a few bytes or even bits. Ethernet becomes rather inefficient when only this type of data is transmitted because the minimum payload of an Ethernet frame is 42 bytes. However, considering the speed of 100 Mbit/s, this is still faster than CAN with 1 Mbit/s. But a lot of performance is lost and processing of Ethernet frames and protocols on transmitting and receiver side requires more performance.
- Looking at robustness, CAN has several mechanisms directly incorporated which makes it well-suited to handle electromagnetic interferences. Erroneous messages are automatically repeated and its residual error probability (probability that a transmission error remains undetected) is several times lower than for Ethernet. In Ethernet, additional software effort is required when the capability for detecting erroneous data shall be improved and messages with erroneous data shall be repeated. Since CAN operates with lower frequencies for data transmission and typically uses smaller microcontrollers, its EMC behavior is better in general.
- Another important advantage of CAN is the possibility to monitor the whole data traffic from any point in the network which is not possible with Ethernet. Here you only see a snapshot of the data which passes by the point of the network which is connected for monitoring purposes. Root cause analysis in case of communication problems becomes much more difficult and typically, it is necessary to have more knowledge to investigate problems.
- The power consumption required for the Ethernet interface of a device is about 3 times higher compared to a CAN interface (considering physical network interface

and required microcontroller interfaces and performance). In case of a stationary machine with a permanent power supply from the electrical network this is not a problem (when not considering heat dissipation or sustainability aspects). However, for battery-driven mobile machines and applications, this is a critical aspect.

- Ethernet is more expensive. Just by comparing costs for the physical interfaces and cables. Running an Industrial Ethernet protocol stack and quite often also a TCP/IP protocol stack in parallel, requires a more performant CPU in order to meet required execution time, response time and real-time demands. Furthermore, more resources are necessary for implementation. This will result is more costs for CPU and memory, especially for frequently used smaller devices with limited functionality. Today, the CAN controller itself is not a cost factor anymore. It is a commodity interface on microcontrollers like a serial interface. In general it can be said that an Ethernet interface is 3 to 5 times the cost of a CAN interface.
- Of course, CAN also has one important disadvantage due to its bus arbitration principle: the maximum extension of a CAN network depends on the transmission speed. At one Mbit/s it is typically less than 40m. Consequently. for larger extensions of the network, e.g. in larger machines, where 100m or 250m are required, the baud rate needs to be lowered to 500 kbit/s or 250 kbit/s.

The conclusion which needs to be drawn from these considerations is that CAN still has its eligibility in many applications, especially when power consumption, price, up-time (MTBF) as well as diagnostic and maintenance capabilities are important requirements.

CAN FD – The booster for CAN

One obstacle which makes the use of CAN in today's automotive and industrial applications difficult, is its limitation to no more than 8 data byes per message. For some applications and functions like electrical drive status and control data in a machine, or wheel rotation information in a car, this requires a split of the transmitted data into two or more messages sent successively. As a consequence, not all data which belongs together arrives at the same time. Also diagnostic data usually consists of longer data packages and therefore needs to be transmitted segmented.

Especially for the car manufacturers, there was a need to increase the bandwidth of the communication networks in the car as more and more electronic functions and ECUs were added. One solution would have been to add more CAN networks and the other solution to use faster networks with higher bandwidth and larger messages. BMW was the first to introduce higher bandwidth by using FlexRay in its cars. Daimler and Audi started using FlexRay somewhat later. However, Opel and General Motors wanted to avoid introducing a new network technology and have been searching together with Bosch for a suitable way to improve CAN in that area.

The result was CAN FD (flexible data-rate) presented in 2012. CAN FD has two essentials improvements: CAN messages can have up to 64 data bytes and bit rates of more than one Mbit/s. In order to keep the advantages and reliability of the arbitration principle but also to be backwards compatible to the "classic" CAN, the transmission speed is only increased after the arbitration phase, when the payload data of a message is transmitted. This principle results in considerably increased bandwidth also for CAN networks with longer line lengths like 250m where the arbitration bit rate still needs to be no more than 250 kbit/s but the payload data then can be transmitted with e.g. two Mbit/s, five Mbit/s or eight Mbit/s.

Today, CAN FD is still not a topic for industrial applications. A main reason for this is that there are no suitable microcontrollers with integrated CAN FD support available yet. However, this is only matter of time, possibly one or two years from now. Meanwhile, all European and American car manufacturers as well as other car manufacturers have plans or have already started introducing CAN FD in their next series of cars. Consequently, the semiconductor manufacturers are already working on new versions of their microcontroller families with CAN FD controllers inside instead of “classic” CAN controllers. When these will become available, it can be expected that also other applications outside of automotive will start using CAN FD.

CAN and the Internet of Things

A big buzzword these days is the “Internet of Things” (IoT). At first glance, it is all about “Ethernet/Internet everywhere”. Ethernet from IT down to the sensor level, the integration of the automation networks with the IT networks and Cloud services.

When looking closer at the use-cases of IoT, it is mainly about improving diagnostics, maintenance and management capabilities of assets (machines and systems) and also about interconnecting machines and systems in order to optimize production and making production more flexible. In the automotive world, it is about communication between cars, between cars and infrastructure and between cars and humans.

The nature of data which is transmitted for these purposes is different than the data transmitted for automation and control purposes (e.g. for controlling a machine). As already discussed in this paper, data used for automation and control is typically short data transmitted more frequently whereas IoT related data is detailed diagnostic data consisting of longer data packages transmitted less frequently.

Within the context of Industrie 4.0 and Industrial Internet of Things, two technologies are considered as core communication technologies: TSN-based Ethernet and OPC UA. TSN (Time Sensitive Networking) will extend Standard Ethernet (802.1) with real-time capabilities. In principle (or should we say theoretically?), the real-time can go down below the microsecond but there are still many open questions and unclear aspects. What seems to be clear is that the faster the real-time will become, the more complex and more expensive the technology will be. Therefore, the primary focus is today to use this technology between IT systems and as “backbone” in areas like the factory floor, and not in sensors and actors inside machines. The purpose of the second technology - OPC UA – is to have a standardized protocol for communicating data between senders and receivers. Of course, there are already many other protocols available for this purpose, but their nature is that sender and receiver already have to know the meaning of the data. There is no information provided by the protocol about the nature and meaning of the data. Hence OPC UA also provides semantics for each data point. This is a strong requirement when data must be exchanged between automation systems and IT systems.

As also already discussed, there is a tradeoff between the use of CAN or Ethernet for automation and control purposes inside a machine. For many applications, CAN remains the preferred solution because of efficiency, costs, reliability, performance, and maintainability. However, related to IoT and the additional data which needs to be gathered from sensors and actuators inside the machine, CAN would not be a suitable solution because of its limited bandwidth with up to 8 data bytes in one message. CAN FD will transmit up to 64 data bytes in one message even at a much higher bit rate. This will allow for collecting the necessary additional IoT-related data from the sensors and actuators inside the machine in parallel to the automation and control data. It will also make it possible to transmit the OPC UA protocol in case devices inside a machine provide OPC UA services.

Therefore, CAN FD is also the perfect sub-network for IoT-related applications, because inside the machine, it will be able to handle automation and control data in parallel to IoT-related data, resulting in overall cost savings.

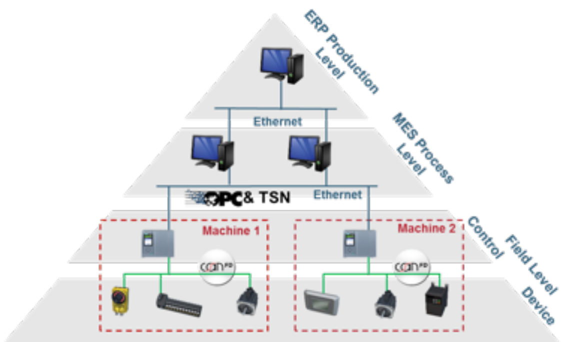

Figure 5: IoT architecture with CAN FD

The future of CAN / CAN FD in automotive applications

Today, CAN is very well established, not only in passenger cars but also more or less in everything moving. Due to higher demands for new functions inside a car, like driving assistance and security, camera systems providing surround views of the car or infotainment, the Ethernet technology also slowly started entering the cars already at the end of the last decade.

For Ethernet, there are specific requirements for use inside cars. First of all, regular CAT5 or CAT6 cables are not appropriate for use inside cars. CAT5/6 cables are too thick and stiff. Therefore it was necessary to develop a new way of transmitting data at 100 Mbit/s in cars. The solution was BroadR-Reach. Identical to Fast Ethernet BroadR-Reach is a point-to-point connection providing a mechanism of 100 Mbit/s full-duplex transmission by using only two wires. Consequently, Ethernet wiring inside the car is similar to CAN wiring. For the mentioned application areas in which Ethernet is already used or is going to be used, Ethernet also needs to have a specific feature for ensuring that data arrives in time at the receivers. Initially intended for the infotainment or camera systems, the AVB standard (Audio Video Bridging) was used consisting of time synchronization based on the IEEE1588 technology and defined bandwidth reservation for data streams. For new application areas like driving assistance and security, the focus is on TSN (Time Sensitive Networking) as this standard is more general and will be part of 802.1.

Will Ethernet replace CAN in cars? - Definitely not within the next 10 or even more years. A major aspect for cars are costs. Today, the car manufacturers consider an Ethernet interface based on BroadR-Reach to be about six times more expensive than a CAN interface. Another cost factor is the fact, that Ethernet is a point-to-point connection requiring switches. This restricts the topology and since switches are only available in a certain granularity, unused ports are an unnecessary cost. It is also not acceptable to use more powerful microcontrollers with MII interface for simple ECUs like those used in doors.

However, the way that data is communicated inside cars has also started to change. In recent years, new mechanisms have been introduced in the AUTOSAR standard like Intelligent PDU Multiplexing (IPDU), Secure on-board communication (SECOC) or E2E (End-to-end protection profiles). These mechanisms require more data to be transmitted than before. In addition, there are authentication mechanisms for ensuring that safety-critical data is only accepted by receivers when the data comes from a known and authorized sender.

Therefore, CAN FD comes at the right time. It will be the bridge between the “classic” CAN world and Ethernet. It is the most cost-effective solution for many use cases inside the electrical car infrastructure. And there is one more very important aspect: CAN is known to everyone and there is a lot of expertise at the car manufacturers. Ethernet is new and there are usually only a few people with the required knowledge. It will take time until everyone has the same knowledge level for Ethernet as for CAN.

All German car manufacturers will introduce CAN FD in their next generations, but CAN will not yet be completely replaced by CAN FD. This will take some time. Ethernet will only be used where the performance and capabilities are absolutely required.

The future of CAN / CAN FD in non-automotive applications

Its cost structure, flexibility, ease of use and low power consumption still makes CAN the best choice for the automation and control networks in many different application areas. Especially when the network extension is limited like in smaller machines, for extensions and sub-systems to larger machines or in all the various mobile applications, and especially when they are battery powered. With CAN FD there is four times or even higher bandwidth available at shorter transmission times, and the network length can also be extended. With the longer CAN FD data packages, CAN also enables the connection of IoT applications down to the sensor level.

Summary

Form today's point-of-view, the assumption that Ethernet will become the only communication system used for automation and control purposes is definitely wrong. There are requirements Ethernet will not meet and the price level and ease-of-use of CAN remains unbeatable. CAN FD comes exactly at the right time to make CAN fit for the future meeting demands for higher bandwidth and performance by keeping the ease-of-use, reliability, flexibility and cost level.

Christian Schlegel
HMS Technology Center Ravensburg GmbH
Helmut-Vetter-Straße 2,
88213 Ravensburg
+49-751-56146-121
+49-751-56146-29
chsc@hms-networks.de
http://www.hms-networks.com