Basic principles of networking

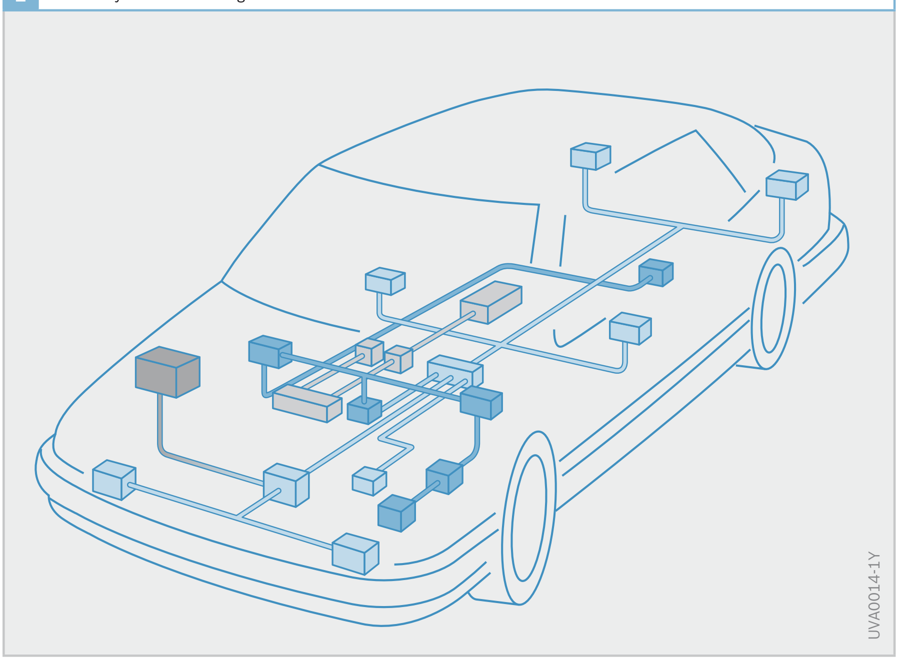
With the tremendous speed at which computer technology is advancing, the number of electronic systems in use is increasing more and more. This growth is also continuing in automotive engineering. However, this also means that the complexity of an overall system (the vehicle in this case) is on the increase. Individual systems such as engine management have been improved over the last few years. However, innovations are mainly achieved by means of interaction between several individual systems. The individual components need to be networked so that the multitude of information that is managed by the individual systems can also be used elsewhere throughout the system. Different communication systems are used depending on requirements (e.g. transmission reliability, fault tolerances, costs).

Network topology

A network is understood to be a system in which a group of elements can exchange information via a transportation medium. If the elements are visualized as nodes and the communication relationships as lines, a picture is created of a network where many nodes are related to several other nodes. The nodes in a communication network are also often referred to as network subscribers or stations.

In motor vehicles, complex control units such as those for the engine management system (Motronic or electronic diesel control, EDC), the electronic stability program (ESP), the transmission control system or the door modules can be network subscribers (Fig. 1). However, a sensor with a conditioning circuit that merely prepares and digitizes a measured value can also act as a network subscriber and make the measured signals available to other net-

1 Vehicle system networking



Robert Bosch GmbH (ed.), *Bosch Automotive Electrics and Automotive Electronics,* Bosch Professional Automotive Information, DOI 10.1007/978-3-658-01784-2_2,

work subscribers. The transport medium via which the communication takes place is referred to as a bus or a data bus.

A network topology is understood to be the structure consisting of network nodes and connections. This merely shows which nodes are interconnected, but does not depict underlying details such as the length of the connection. Every network subscriber must have at least one connection to another network subscriber in order to participate in network communication. Different requirements are made of the network topology for a variety of communication network applications, while the topology determines some of the characteristics of the overall network. All network topologies are based on the following four basic topologies
- Bus topology
- Star topology
- Ring topology and
- Mesh topology

Other structures (hybrid topologies) can be created by combining these basic topologies.

Bus topology

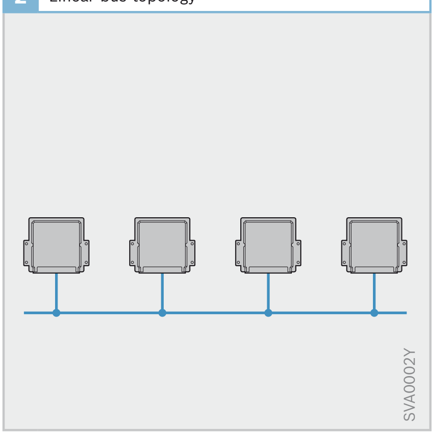
This network topology is also referred to as a linear bus. The core element of a bus topology is a single cable to which all nodes are connected via short connecting cables (Fig. 2). This topology makes it extremely easy to add other subscribers to the network. Information is transmitted by the individual bus subscribers in the form of so-called messages and distributed over the entire bus.

Nodes transmit and receive messages. If a node fails, the data that is expected from this node is no longer available to the other nodes on the network. However, the remaining nodes can continue to exchange information. However, a network with a bus topology fails completely if the main line is defective (due to a cable break, for example).

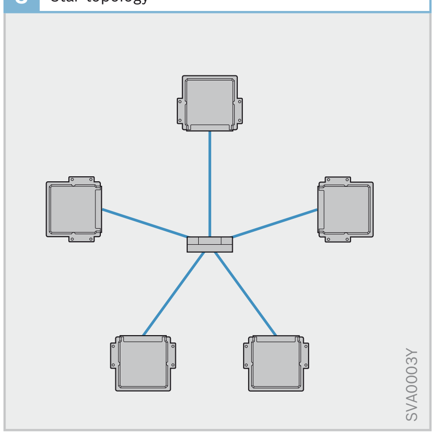
Star topology

The star topology consists of a main node (repeater, hub) to which all other nodes are coupled via a single connection (Fig. 3). A network with this topology is therefore easy to extend if free capacity is available (connections, cables).

2 Linear bus topology



3 Star topology



In star topologies, data is exchanged between the individual node connections and the main node, whereby a distinction is made between active and passive star topologies. In active star topologies, the main node contains a computer that processes and relays information. The performance capability of a network is essentially determined by the performance capability of this computer. However, the main node does not have to have special control intelligence. In passive star systems, it merely connects the bus lines of the network subscribers together.

The following applies to active and passive stars: if a network subscriber fails or a connecting line to the main node is defective, the rest of the network continues to operate. However, if the main node fails the entire network is disabled.

In the automotive area, star structures are under discussion for safety and security systems such as brakes and steering. In this case, the risk of a complete network failure is prevented by designing the main node to be physically redundant. This means that several main nodes are used to which the nodes whose information is needed for safe operation of the vehicle can be connected in parallel.

Ring topology

In the ring topology, each node is connected to its two neighbors. This creates a closed ring (Fig. 4). A distinction can be made between single rings and double rings.

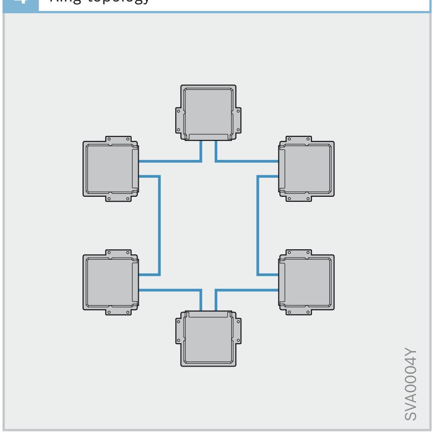
In a single ring, data transfers are unidirectional from one station to the next. The data is checked when it is received. If the data is not intended for this station it is repeated (repeater function), boosted and relayed to the next station. The data that is being transferred is therefore relayed from one station to the next in the ring until it has reached its destination or arrives back at its point of origin, where it will then be discarded.

As a soon as a station in a single ring fails, the data transfer is interrupted and the network breaks down completely.

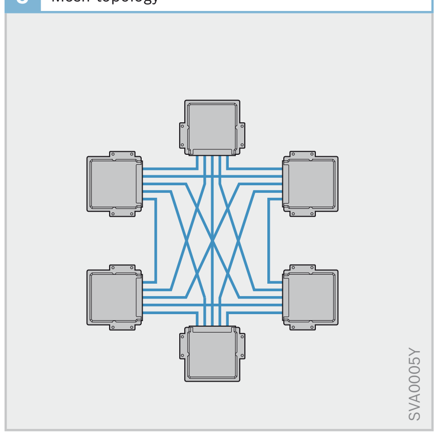
Rings can also be set up in the form of a double ring (e.g. FTTI), in which the transfer of data is bidirectional. In this topology, the failure of a station or a connection between two stations can be overcome, since all data is still transferred to all operational stations in the ring.

However, if several stations or connections fail, the possibility of a malfunction cannot be ruled out.

4 Ring topology



5 Mesh topology



Mesh topology

In a mesh topology, each node is connected to one or more other nodes (Fig. 5). In a fully meshed network, each node is connected to every other node.

If a mode or connection fails it is possible for the data to be rerouted. This type of network therefore has a high degree of system stability. However, the cost of networking and transporting the message is high.

Radio networks form a type of mesh topology, since the transmissions from each station are received by every other station that is within range.

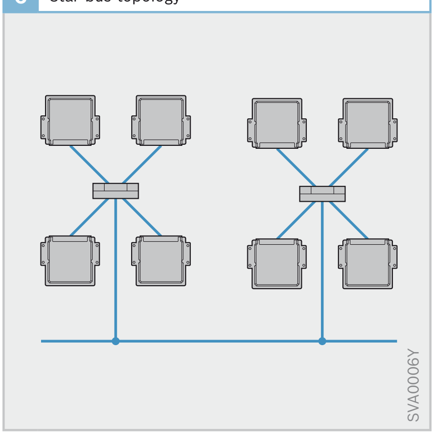
A mesh topology is bus-like as far as exchanging messages is concerned, and star-like regarding data transfers, since every station receives all transmissions from every other station, but connection failures can be overcome.

Hybrid topologies

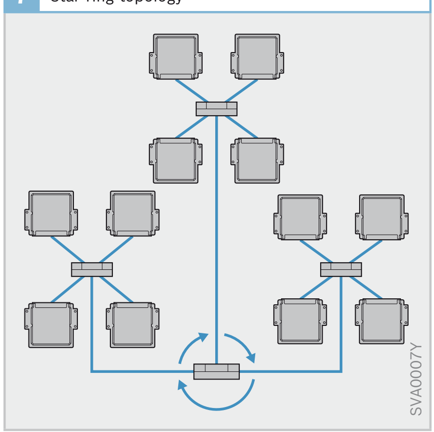
Hybrid topologies are a combination of different network topologies. Examples of such combination are:

- Star bus topology: the hubs of several star networks are interconnected as a linear bus (Fig. 6).
- Star ring topology: the hubs of several star networks are connected to the main hub (Fig. 7). The hubs of the star network are connected in the form of a ring in this main hub.

6 Star bus topology



7 Star ring topology



Network organization

Addressing

In order to make it possible to transmit messages via a network and evaluate the contents thereof, the useful data (payload) that is transmitted is also accompanied by data transfer information. This can be explicitly contained within the transmission or implicitly defined using preset values. Addressing represents important information for data transfer information. It is needed in order for a message to be sent to the correct recipient. There are different ways of doing this.

Subscriber-oriented method

The data is exchanged on the basis of node addresses. The message sent by the transmitter contains the data to be transmitted and also the destination node address (Fig. 8a). All receivers compare the transmitter receiver address to their own address, and only the receiver with the correct address evaluates the message.

The majority of conventional communication systems (such as Ethernet) operate using the subscriber addressing principle.

8 Addressing modes

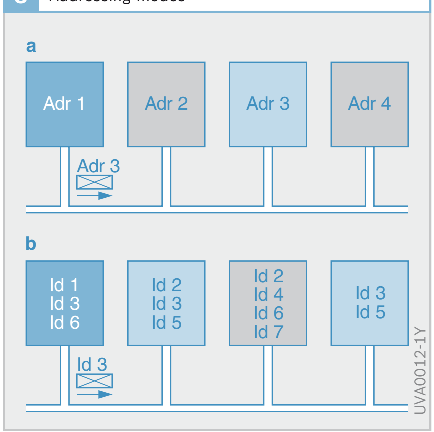


Fig. 8
a Subscriber-oriented method
b Message-oriented method

Message-oriented method

In this method it is not the receiver node that is addressed, but the message itself (Fig. 8b). Depending on the content of the message, it is identified by a message identifier that has been predefined for this message type. In this method, the transmitter does not need to know anything about the destination of the message, since each individual receiver node decides whether or not to process the message. Of course, the message can be received and evaluated by several nodes.

Transmission-oriented method

Transmission characteristics can also be used to identify a message. If a message is always transmitted within a defined time window, it can be identified on the basis of this position. By way of a safeguard, this addressing is often combined with message or subscriber-oriented addressing.

Bus access method

A node must access the bus in order to transmit a message. In the bus access method a distinction is made between

- Predictable methods in which the bus access is determined by certain time-dependent network characteristics, whereby only one node can transmit at a time and
- Random methods whereby any node can attempt to transmit data if the bus appears to be free

In the predictable method the bus access right is determined before bus access. It can thereby be ensured that only one subscriber is using the bus at a time. Access collisions because of simultaneous bus usage will be prevented if all subscribers use this method.

In the random method, the nodes can simultaneously attempt to use the bus as soon as it appears to be free. The timing of the bus access is therefore random. There is a risk of transmission collisions using this method, which will require attention. This can be dealt with by repeating transmissions after a collision has been detected (e.g. Ethernet), by giving the transmissions different codings (CDMA), controlling communication via a master or prioritizing message types or transmitters.

Time division multiple access (TDMA)

TDMA is a deterministic (predictive) access method. In this case each node is assigned a time window in which it is allowed to transmit (a priori). A fixed schedule is therefore required for the network. There is not usually a main communication subscriber controlling the communication procedure. However, concepts exist in which in which it is possible to switch between different schedules if necessary. The internal clocks of the different stations must run extremely synchronously with TDMA, since the transmit windows have to be adhered to with extreme precision.

Master-slave

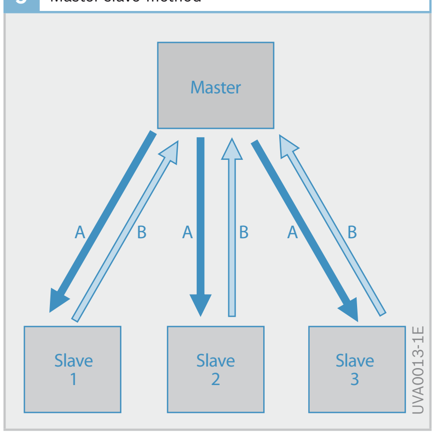
In the master-slave system, one node on the network operates as the master. This node determines the communication frequency by interrogating its subordinate nodes (slaves). A slave only replies if it is spoken to by the master (Fig. 9). However, some master-slave protocols allow a slave to contact a master in order to transmit a message (e.g. transmit information about the position of the power-window unit to the door module).

Multimaster

In a Multimaster network, several nodes can access the transport medium independently without the assistance of another node. Bus access is uncontrolled. Every node can access the bus and transmit a message if the bus appears to be free. This means that each node is its own master, and that any node can start a message transfer with equal status. However, this also means that collision detection and handling methods have to be in place. For example, this may be in the form of a decision-making phase with prioritization or delayed transmission repeats. The use of priority control prevents a bus conflict if several nodes wish to use the bus at the same time, since the network node that has high priority or wishes to transmit a message with high priority forces its way through in the event of a conflict and transmits its message first. Normal message transmission resumes when the line is free again.

The Multimaster architecture has a positive effect on the availability of the system, since no individual node is in control of communication whose failure would lead to total communication breakdown.

9 Master-slave method



OSI reference model

Network protocols are usually defined in layers, which combine properties and tasks. The properties of the deeper-lying layers are assumed in the next level up. This has the advantage that individual layers are exchangeable, provided that the interfaces that are provided between the layers remain unchanged.

The ISO OSI reference model (Open Systems Interconnection) provides a basis for describing and comparing many communication protocols. This was developed by the ISO (International Standardization Organization) and led to the adoption of international standards by ISO and IEEE (Institute of Electrical and Electronic Engineers).

In the OSI model, data communication systems are depicted in different layers (Fig. 10). The complex task of data communication is distributed among clearly arranged functional areas (layers). Not all of the layers in the OSI model are needed in a simple communication system. Layers can also be combined for many applications. Network protocols in the automotive area are often divided up into

- Physical layer
- Communication layer and
- Application layer (user layer)

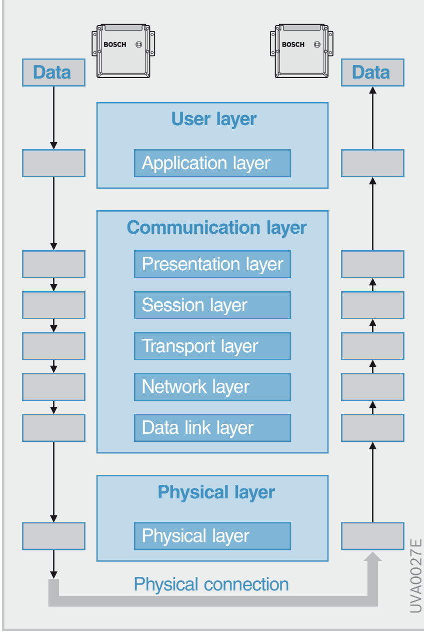
Physical layer

The electrical and procedural parameters of the physical connection between the network subscribers is defined in the physical layer.

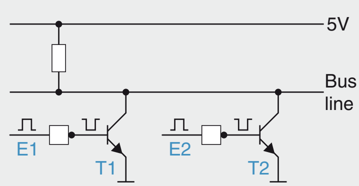
Signal level

In digital technology, data is represented by sequences of the two binary statuses, 0 and 1. In order to transmit the data on a bus, these statuses must be represented on the transmission agent. It is particularly important to avoid short-circuits on the

10 OSI reference model



11 Creating dominant and recessive levels



Control E	Status T	Bus level
0	conducts	0 V
1	blocks	5 V

E1	E2	Bus level	
0	0	0 V	dominant
0	1	0 V	dominant
1	0	0 V	dominant
1	1	5 V	recessive

SVA0008Y

bus when one node is transmitting a status of 1 and another is transmitting a status of 0.

The binary statuses can be depicted in many different ways. The serial interface of the PC, for example, uses +12 V and -12 V, and CAN-B uses voltages of 0 V and 5 V. The voltages of the serial interface are unsuitable for a bus, since short-circuits can occur if several subscribers wish to transmit conflicting binary statuses simultaneously.

If the coding allows one level to overwrite another, the overwriting level is referred to as dominant, and the subordinate level as recessive.

It is also possible to depict dominant and recessive levels using visual media. A status of 1 (recessive) then corresponds to e.g. dark, and a status of 0 (dominant) corresponds to light. In an optical fiber, an individual node can override all of the others by feeding light into the conductor.

Bit stream

The application information cannot usually be transmitted directly. In order to make transmission possible, the information is first incorporated as a payload in the frame of a message that contains information to be transmitted. Since all protocols have been developed in accordance with different requirements, the frame format differs from protocol to protocol.

The frame needs to converted into a bit stream to actually transmit the information, i.e. a sequence of bits that can be transmitted via the transport medium as physical states.

Example of the UART interface

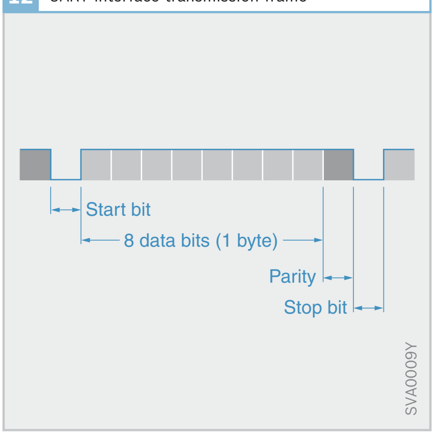
The microcontrollers that are used in control units have a simple interface (UART, Universal Asynchronous Receiver/Transmitter) on the chip, via which they can communicate with the outside world (e.g. with a PC). The essential features of a data transmission are read out via this interface.

If no data is being exchanged, the bus level is 5 V (microcontroller operating voltage, Fig. 11). When the start bit is transmitted (dominant level), the other station connected to the bus (receiver) is notified that a data transfer is starting (Fig. 12). The length of the start bit determines a bit time that represents the basis for the entire data transfer. Every subsequent data bit has the same length. The reciprocal of this time corresponds to the data transfer rate, i.e. the number of bits that can be transmitted in one second in a continuous data stream. All participating stations must be set to the same data transfer rate.

After the start bit has been received, the transmission of an 8-bit data word commences (1 byte) with the lowest significant bit (LSB, Low Significant Bit). The receiver that has synchronized itself to the start bit scans the data bus between each data bit and therefore assembles the transferred data byte.

The eighth data bit is followed by the parity bit. This bit indicates whether the number of transmitted ones is odd or even. It therefore allows the receiver to perform a simple check for possible transmission errors. The sequence is completed with the stop bit, which is placed onto the bus

12 UART interface transmission frame



with a dominant level. The next data transfer can then take place.

Communication layer

Control units can only interconnect and exchange data if they speak the same "language". This language determines the rules that are used to exchange data between the individual network subscribers.

The communication layer accepts data from the application layer, prepares it for transmission and forwards it to the physical layer.

The essential features of this protocol layer are:

- Message frame format
- Bus access control
- Message addressing
- Detection and handling of collisions
- Network node synchronization
- Checksum calculation

Application layer

The application layer consists of the application that processes and provides the information. The application layer is the only protocol layer to be affected by user or sensor input.

Control mechanisms

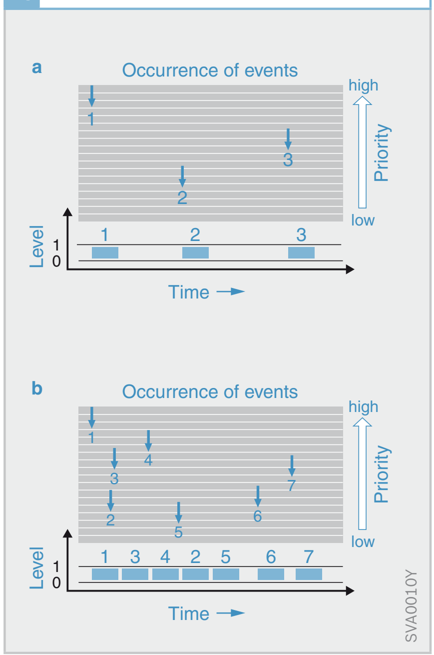
Event control

In an event-driven bus system, messages are transmitted as soon as an event that triggers the transmission of a message has occurred (Fig. 13a). Examples of such events are:

- Pressing a button on the air conditioning system control panel
- Operating the hazard warning flasher switch
- Incoming message that requires a reaction (e.g. information from rpm sensor to speedometer needle motor)
- Expiration of a fixed time period (time frame, e.g. 100 ms), after which messages are transmitted cyclically

Since the stations are not synchronized with each other, situations where several stations wish to access the bus simultaneously are unavoidable. In order to allow

13 Event control



a message to be transmitted without falsification, only one station at a time can transmit data on the bus. Collision avoidance mechanisms are available for preventing or solving bus conflicts.

If a node wishes to transmit a message whilst the bus is occupied, the transmission is delayed (Fig. 13b). A station that is ready to transmit must then wait until the transmission that is currently in progress has been completed.

Since bus access is subsequently renegotiated, the transmission may be delayed yet again. These delays become problematic if the bus becomes overloaded by a large number of network subscribers that wish to transmit messages. In this case messages may be lost if the transmitter abandons the transmission due to excessive delays.

Event-driven bus systems are suitable for reacting to asynchronous (unforeseen) events as quickly as possible. In an ideal case, they reduce the delay between the occurrence of the event and the message transmission (latency time) compared to time-driven systems. However, the latency time can vary considerably depending on the network loading.

Advantages

- High level of flexibility and capability of retrofitting new nodes in the network
- Good response time to asynchronous external events
- Bus usage depending on event frequency in line with requirements
- No network loading by unused events, since only events that have actually occurred trigger a transmission

Disadvantages

- Static bus occupancy, non-deterministic (i.e. not possible to prove that a message was transmitted at the right time)

Timer control

In the most recent developments in dynamic driving systems such as brakes and steering, an increasing number of mechanical and hydraulic components are being replaced with electronic systems (x-by-wire). Mechanical connections such as the steering column are becoming superfluous, and the functionality thereof is being taken over by sensors and actuators. The reliability, safety and failure tolerance requirements of these systems are extremely high. This means:

- Messages must be received on time
- The latency time of critical messages must be extremely small
- The system must have a redundant design
- The failure of a node must affect the rest of the system as little as possible and
- It must be possible to achieve a safe operating status from any fault situation

X-by-wire systems require close networking by the various components. The external increase in complexity places new demands on the safety, failure tolerance and availability of the communication system. The demands that are made of the electronic and network architecture therefore also increase. A reliable, fault-tolerant network architecture is required so that data is transmitted with guaranteed transmission characteristics, and electronic system malfunctions are handled in the most efficient way.

System architectures for real-time applications meet these requirements because their behavior is predictable and verifiable because of the way in which they are constructed. In these protocols, time windows within which a node is permitted to transmit are assigned to the control units in the communication network (nodes) during network planning (Fig. 14). In order to comply with the time window, the nodes must be synchronized as precisely as possible.

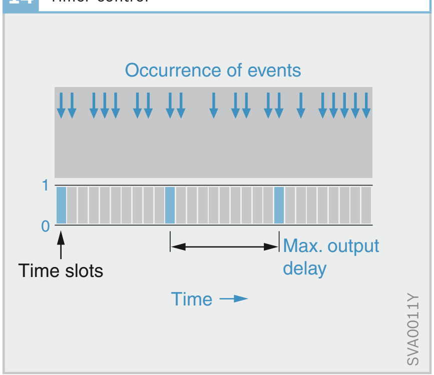
All transmissions are processed sequentially in accordance with the network planning (without collisions). Once each node has transmitted its message, the cycle restarts with the first transmitter. This makes it possible to determine how chronologically up-to-date the data is at any time. Since missing messages are detected immediately, time-triggered concepts are more reliable than event-driven systems.

If a fast data rate is required in a time-triggered system, the time delay between the occurrence of an event and the transmission of the data can be so small that the system complies with strict real-time requirements.

The bus can be protected from unauthorized access by a bus guardian. The bus guardian prevents a defective node from interfering with network communication by transmitting messages outside the relevant transmit window.

These characteristics make it possible to create redundant, fault tolerant systems in which transmission errors can be remedied and faults in the network can picked up by network nodes that can provide the functionality without errors.

14 Timer control



Advantages
- Deterministic system
- Punctual data transmission
- Reliable detection and isolation of defective network nodes

Disadvantages
- Overall system must be planned for distributed developments
- Capacity for expanding the communication system must be planned in
- Good response time to asynchronous external events

Composability
If a communication system allows independently developed subsystems to be integrated in an overall system, it is said to support composability. An important criterion when doing this is that the properties that have been assured for the functionality of a subsystem are not adversely affected by adding other subsystems. If this has been ensured, the checking of system functionality is restricted to subsystem checking that can be carried out by the constructor of the subsystem.

If a communication system supports composability, changes can be made to a control unit without affecting the functionality of other control units. It is therefore not necessary to recheck the entire system after integrating a modified control unit – it is sufficient to check that the individual subsystems are operating reliably. Composability therefore reduces the time and cost of integrating new subsystems. This is the only way to increase the complexity of the electronics in the vehicle.

Overview of bus systems used in vehicles

	CAN-C high-speed CAN	CAN-B low-speed CAN	LIN	TTP
Definition	Controller area network	Controller area network	Local interconnect network	Time-triggered protocol
Bus type	Conventional bus	Conventional bus	Conventional bus	Conventional and optical bus
Domains	Drivetrain	Comfort/ convenience	Comfort/ convenience	Safety-related networking
Applications	Engine management, transmission control and ABS/ESP networking	Body and comfort and convenience electronics networking	Low-cost expansion of CAN bus for simple applications in the comfort and convenience electronics area	Networking in safety-related environments such as brakes, steering, railway signal boxes or aircraft landing gear
Most frequently used topology	Linear bus	Linear bus	Linear bus	Star topology
Data transfer rate	10 kbit/s to 1Mbit/s	Max. 125 kbit/s	Max. 20 kbit/s	Unspecified, typ. 10 Mbit/s
Max. number of nodes	10	24	16	Unspecified
Control mechanism	Event-driven	Event-driven	Time-driven	Time-driven
Bus lines	Copper conductors (twisted pair)	Copper conductors (twisted pair)	Copper conductor (single wire)	Copper conductors (twisted pair)
Deployment	in all vehicles	in all vehicles	in all vehicles	Premium class vehicles, aircraft, rail control systems
Standard	ISO 1198	ISO 11519-2	LIN consortium	TTA group
SAE classification	Class C	Class B	Class A	Drive-by-wire

	MOST Bus	Bluetooth	Flexray
Definition	Media oriented systems transport	Proprietary name (Danish king)	Proprietary name
Bus type	Optical bus	Wireless	Conventional and optical bus
Domains	Multimedia and Infotainment	Multimedia and Infotainment	Deployment across all domains
Applications	Transmission of control, audio and video information	Data transfers over short distances, e.g. mobile phone integration in the infotainment system	A network system for use in safety-related and simple applications
Most frequently used topology	Ring topology	Network topology (radio)	Star topology
Data transfer rate	Max. 22.5 Mbit/s	Max. 3 Mbit/s (v2.0) Max. 723 kbit/s (v1.2)	Typ. 10 Mbit/s Max. 20 Mbit/s
Max. number of nodes	64	8 active (up to 256 passive)	Theoretically up to 2,048 Max. 22 per passive bus/star
Control mechanism	Time and event-driven	Event-driven	Time and event-driven
Bus lines	Plastic or glass optical waveguides	Electromagnetic radio waves	Copper conductors (twisted pair)
Deployment	Premium class vehicles made by European manufacturers	All vehicles, connection between multimedia equipment and infotainment system	Pilot application
Standard	MOST cooperation	Bluetooth SIG	Flexray consortium
SAE classification	Mobile Media	Wireless	Drive-by-wire

Table 1

http://www.springer.com/978-3-658-01783-5

Bosch Automotive Electrics and Automotive Electronics
Systems and Components, Networking and Hybrid Drive
(Ed.)
2014, IX, 521 p. 595 illus. in color., Softcover
ISBN: 978-3-658-01783-5