Basic Communication Protocols in Vehicle Electronic Systems

Dragan S. Taranović[1], Jasna J. Radulović[2], Saša J. Jovanović[3], Andrija M. Savčić[4]

Abstract – **Today's vehicles are being equipped with a constantly increasing number of different electronic systems. Along with their need for extensive exchange of data and information in order to operate efficiently, the data quantities and speeds are also increasing continuously. In this paper, the analysis of some serial network protocols used in vehicles is presented.**

Keywords – **Vehicle networks, CAN, TTP.**

I. Introduction

Improvement in vehicle driving characteristics, driver and passenger comfort and reduction of vehicle exhaust emission is achieved through application of electronic control systems having a large number of sensors and actuators.

From the aspect of functionality, vehicles contain three parts, which can be independently controlled and which should exchange information based on appropriate standard procedure - protocol:

- control system for power unit, brakes and vehicle dynamics,
- system for signalization and diagnostics,
- system providing comfort and driver and passenger safety (air-conditioning, central lock-up, adjustment of seats and mirrors, lights, navigation, ...).

Classic system for data transfer and vehicle functions control has a star-like structure shown in Fig. 1., in which there is one main control unit (master), marked as MCU, and several local control units, LCU, sensors, S, and actuators, A, connected with the main control unit through local, serial or parallel connection. Processing of sensor signals and actuator commands can be realized in main microprocessor control unit or in appropriate local control system.

Star-like configuration of control system demands that main control system has a high speed processor in order to process all acquired information. Wire installation for connecting the sensors and actuators with control units is highly complicated and very long, and it makes the interaction between different systems harder, making the whole system susceptible to disturbances.

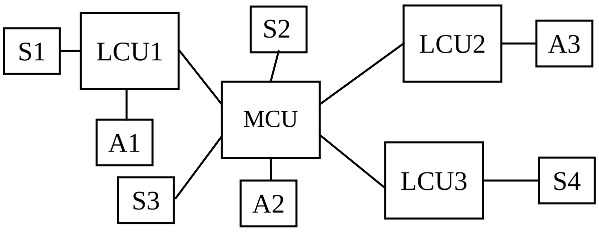

Fig. 1. Star-like configuration of control systems

A better quality of vehicle control is achieved by application of, usually, double-wire network to which all vehicle systems control units, "intelligent" sensors and "intelligent" actuators are attached on and communication is serial. Scheme of such a network is presented in Fig. 2. Main control unit need not to be present in this type of a network.

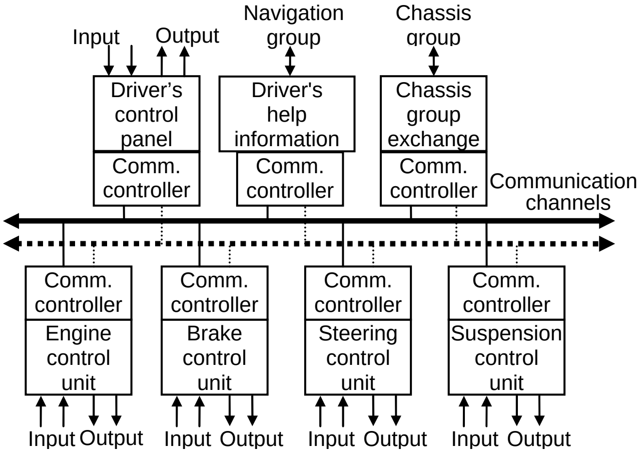

Fig. 2. Serial configuration of control systems

The vehicle, as a complex distributed electronic system, is composed of several network interconnected subsystems called “groups” (clusters) [2]. Each group realizes a separate function. Examples of groups are: electronic systems for drive unit control and vehicle dynamics control, electronic chassis systems (signaling and comfort), multimedial and electronic navigation systems, etc.

[1] Dragan Taranović is with the Faculty of Mechanical Engineering, Sestre Janjić 6, 34000 Kragujevac, Serbia and Montenegro, E-mail: tara@kg.ac.yu

[2] Jasna Radulović is with the Faculty of Mechanical Engineering, Sestre Janjić 6, 34000 Kragujevac, Serbia and Montenegro, E-mail: jasna@kg.ac.yu

[3] Saša Jovanović is with the Zastava Automobiles – R&D Department, Automobile Institute, Trg topolivca 4, 34000 Kragujevac, Serbia and Montenegro, E-mail: piter@ia.kg.ac.yu

[4] Andrija Savčić is with the Zastava Automobiles – R&D Department, Automobile Institute, Trg topolivca 4, 34000 Kragujevac, Serbia and Montenegro, E-mail: piter@ia.kg.ac.yu

II. VEHICLE NETWORKS

The vehicle usually has several networks with different hardware and software realizations, that is, they operate using different protocols [1]. Network and network protocols for automotive applications must have the following basic characteristics:

- high integrability composed with other components and systems of the vehicle as a whole, that is, probability of occurrence of random error should be negligible and should not affect the vehicle functioning,
- functional adjustment, that is, maximal waiting time for a transfer and to transfer information should be short enough in order not to interfere with control,
- configurability of the network, that is, the network can be easily broadened and modified,
- error resistance, that is, communication has to be reconstructed when error is debugged, while the existence of the spare channel is advisable and, sometimes, necessary,
- minimal number of interconnections, that is, every additional connection/connector increases the probability of error occurrence,
- dimensionally and functionally adjusted connectors,
- electro-magnetic compatibility, ,
- environment resistance – resistance against temperature, humidity, vibration, dust, fuel drops, oils, lubricants, etc,
- low price.

The vehicle network should provide a connection to other computer systems outside the vehicle (navigation, diagnostics, vehicle ecological parameters control...). That is why it must also satisfy general standards of computer equipment connections given by ISO7498 which defines the Open System Interconnection (OSI) with seven basic interconnection layers.

Three layers of network interconnection are enough for the use in vehicles: layer 1- physical layer, layer 2- data link layer and layer 7- application layer, although there are protocols defining all seven interconnection layers (SAEJ1939).

Application of networks in control systems of vehicle systems has induced a need for appropriate protocol for information exchange between network nodes. Several types of protocols has appeared because producers reacted differently and named the protocols as follows: Volkswagen - ABUS protocol, Renault & PSA - VAN protocol, Toyota - BEAN protocol, General motors - J1850VPW protocol, Ford - J1850PWM protocol...

Society of Automotive Engineers (SAE) has divided vehicle networks into three classes, depending on information transfer rate through network: A, B and C. Fast development of multimedia applications (Internet, digital TV), as well as, so called "control by wire", X-By-Wire (e.g. vehicle control system), demand higher data transfer rates, resulting in formation of yet not standardized class D. Basic characteristics and purpose of classes mentioned are given in Table I.

Protocols can be divided into event-triggered and time-triggered protocols, according to manner in which they produce information.

TABLE I

Network class	Transfer rate	Application
A	<10kb/s Small speeds	Driver and passenger comfort; mirror and seat adjustment, opening the trunk, central lock-up
B	10-125kb/s Medium speeds	Vehicle instruments, vehicle speed, exhaust emission data
C	125kb/s-1Mb/s High speeds	Real time control of engine functions, vehicle dynamics, braking system
D	>1Mb/s	Real time control of systems responsible for driver and passenger safety; multimedia applications

III. EVENT TRIGGERED PROTOCOL

Event triggered protocols operate on the principle of generating information from a network node when there has been a change in information sent by that node. This type of protocols is usually implemented by serial network.

Event triggered protocols are: CAN (Controller Area Network), LIN (Local Interconnection Network), Byteflight, etc.

Bosch, with it's CAN protocol, formed in 1985, and given the actual form in 1991, has become a leading manufacturer in the area of network protocols for automotive application. Bosch has made it's protocol opened for all users, making it accepted very soon and a basis for ISO11898 and SAE standards for vehicles and also for industry and other areas.

CAN network is made of nodes that have a structure shown in Fig. 3., according to the OSI interconnection model.

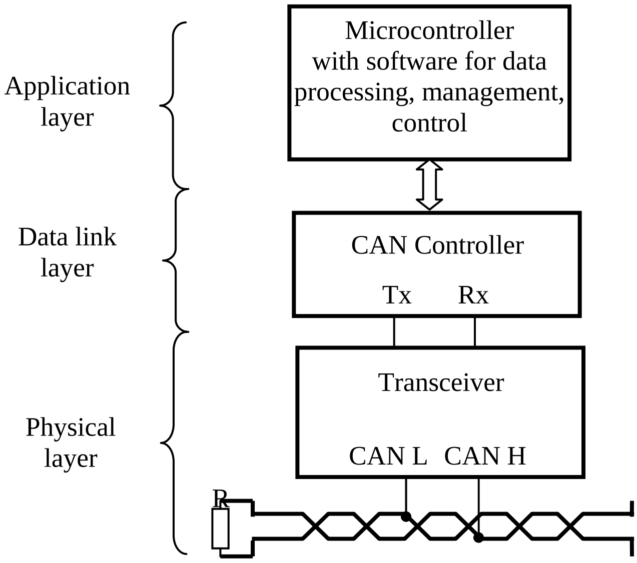

Fig. 3. Structure of CAN network node

The principle of operation of the network based on CAN protocol is based on the fact that only one network node sends the message at one moment and the message is received by all network nodes, including the one that sends it. The network should have the following features:

- Every message sent by a network has appropriate priority.
- Lower priority messages are guarantied to be sent after a certain time delay.
- The network should have flexibility in adding new nodes to the network.
- Synchronized receiving and transmission between all network nodes.
- Changeable format of data sent through the network.
- Several master nodes in the network.
- Error detection and signalization.
- Automatic repetition of inaccurate messages as soon as the network is available.
- Distinction between temporary and permanent errors in nodes and autonomous disconnection of inaccurate nodes.

Physical layer of CAN protocol consists of a line and a transceiver with appropriate connector.

The line as a medium is a double-wire cable with spanned conductors or with the shield. There are types of lines with only one conductor, where the mass is used as a second conductor. The line must be closed with characteristic resistance near to 120 Ω.

Transceiver is a part of the network node representing an electronic circuit which is directly connected to the line and which enables a two-way communication. It separates other electronic node parts from the line. If the line is long, it electrically separates the nodes from the line on which different disturbances can occur due to long length. Transceiver conducts a voltage processing of line data during the data receiving and sets the line in appropriate state during the data delivery. CAN communication go through the bits which can be dominant (logical 1) or recessive (logical 0).

Basic assumption for uninterrupted CAN protocol based network operation is that all network nodes are set to send and receive information at the same rate, that is, they must be synchronized.

Network message should have specific format and consist of segments. Basic message segments are shown in Fig. 5. The message begins with a dominant bit, representing the start of the message (SOF), then there are: the address or the identifier of a node emitting the message (IDENTIFIER), control part of a message (data length and type) (CONTROL), data, key for message accuracy checking and at the end (CRC), there is a confirmation on the message accuracy from all network nodes (ACK). The message is followed by a period of network inactivity in order to prepare the network for processing the next message (EOF).

Two basic CAN protocol types are defined in relation to identifier's length: Standard CAN protocol with the identifier 11 bits long, through CAN protocol 2.0A and Extended CAN protocol with the identifier 29 bits long, through CAN protocol 2.0B. Within the same network, nodes that support CAN 2.0B protocol can exchange information with nodes supporting CAN 2.0A protocol, while reverse exchange is not possible.

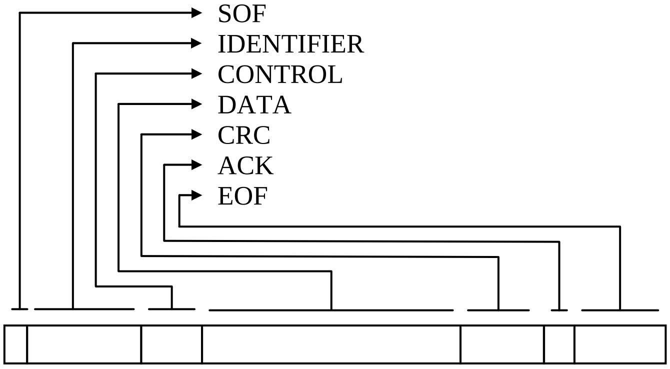

Fig. 4. Basic segments of CAN message

Message priority is determined by identifier, because every node has unique identifier. If two or more nodes start emitting the message at the moment when the line is available, a priority is given to the node which has a dominant bit at the difference in identifiers. Lower priority nodes stop emitting a message and transfer to receiving regime and the emitting of a new message waits until the line is available again.

The length of data in the message can vary between 0 to 8 bytes and it is defined by control bits sent before data.

Accuracy of received message is controlled by a sequence of 15+1 bits generated by a message sender and based on algorithm defined by standard, and it represents a segment of the message (CRC - Cyclic Redundancy Check).

Errors that occur during data transfer can have different shapes and origins, and they are detected by hardware and software processing of received signal. Basic hardware error detection is obtained by a node emitting a message, by comparison of the sent bit with the one read from the line. The next level in error detection is examination if there are 6 consecutive bits of the same type, because it is forbidden by a protocol. Error is also signalized in the case when there is an illegal bit at the end of constant length sequences (identifier, CRC, the end of the message...). Software error detection is obtained by checking of the message with the help of CRC sequence.

The line is available when at least 10 consecutive recessive bits are detected, 7 of which represent the end of the previous message and 3 or more represent a delay for next message preparation.

CAN protocol application layer is specific to every network node. At the application layer, network control units perform processing of data received by a network, as well as data received from the transducer locally connected to the control unit, and generate data necessary for sent by a network.

IV. TIME TRIGGERED PROTOCOL

Messages from TTP (Time Triggered Protocol) are sent in time intervals set in advance [6]. Total time by which all messages should occur is divided into intervals, sections, and at every node, a section is determined to which the message should be sent. All network nodes have access to the network and can read all other messages, including their own. Due to time definition of moment at which the message is sent, the network must be timely synchronized, which is usually achieved by a special node that sends synchronization pulses to all nodes.

Time triggered protocols are: TTP/C, FlexRay®, etc. Their basic application is anticipated for systems critical from the aspect of safety, like “X-by-wire” systems, where the existing mechanical systems are substituted by mechatronic systems. Mechatronic systems have no reserve mechanical systems to take over the control in the case of electronic system failure, so it is necessary that electronic systems with their communication channels are made in such a way to be immune against any kind of error.

In order that time triggered protocol is able to function, it is necessary that each node should have separate TTP controller for protocol control, independent on the process that has been controlled at the node [6]. TTP controller has a message descriptor list (MEDL). Message description contains time position of the section to which the message is sent, the length of the message and serial number of the message sent, because messages with different contents can be sent to one section, but according to order set in advance. The scheme of network node access is shown in Fig. 6.

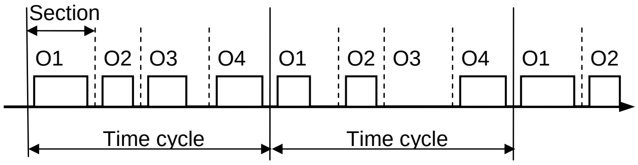

Fig. 5. Network node access at time triggered protocol

The message has a format set in advance, Fig. 7. The header contains information on sent data type, because there are two basic information types sent at DATA area: normal data used for sending information necessary for control and initialization data used for sending information on the state of TTP controller and network synchronization.

At time triggered protocol, the most dangerous error is so called “Bubbling idiot failure”, which manifests through uncontrollable sending of messages from the node outside defined sections. The problem is solved by setting a node that monitors the network traffic (BUS Guardian) and occasionally can exclude the node from the network or perform the re-initialization of the network.

Determinancy of the time triggered protocol creates a great reliability in control systems, with even network loading, but it aggravates inclusion of new nodes into the network. Addition of new nodes requires a total reconfiguration of the entire network.

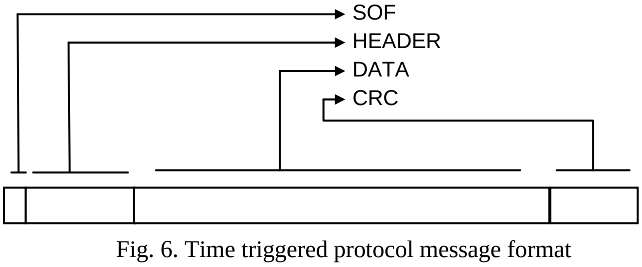

Fig. 6. Time triggered protocol message format

V. Time triggered CAN protocol

Standard CAN protocol provide access to the line for all network nodes, but because of priority principle, some messages can be held back and, after a long enough time period, when they finally can be emitted, they became useless. Information loss is not desirable, especially in systems where "X-by-Wire" control is wanted. To avoid information loss, a time synchronization of network nodes and assignment of time interval when a node should emit it's message are proposed [5]. Time synchronization does not change a basic conception of CAN protocol and such a network is referred to as a Time Triggered CAN - TTCAN.

TTCAN divides time access to the network in such a way that, in certain time intervals, only some nodes - important for reliable control system functioning, have access to the line, while in other intervals, standard CAN protocol regimes are applied. In order to enable the system to function in this way, it is necessary for every node to have local clock which is occasionally synchronized with the network clock.

VI. Conclusions

Standard network protocols with serial communications enable great flexibility in introduction of new vehicle systems and in improvement of existing control units characteristics, but it is necessary to work on their development in order to enhance data transfer speed and reliability in data transfer, which are necessary for "drive-by-wire" principle.

Acknowledgement

The paper is the result of the research within the project financed by the Ministry of science and environmental protection of the Republic of Serbia.

References

[1] Rushby J.: "Bus architectures for safety-critical embedded systems", *EMSOFT 2001: Proceedings of the First Workshop on Embedded Software*, Volume 2211 of Lecture Notes in Computer Science, str.306-323, Lake Tahoe, SAD, Springer-Verlag, 2001

[2] Leen G., Heffernan D., Dunne A.: "Digital Networks in the Automotive Vehicle," *IEE Computer and Control Eng. J.*, p.p. 257-266, 1999

[3] Taranovic D., Grujovic A.: "Standardization of network protocols in vehicle electric interconnection", *Mobility & Vehicle Mechanics*, Volume 28, Kragujevac, 2002

[4] Robert Bosch Gmbh: *"CAN Specification Version 2.0"*, Stuttgart, 1991

[5] Hartwich F., Müller B., Führer T., Hugel R.: "CAN Network with Time-triggered Communication", *7nd International CAN Conference, CAN in Automation (CiA)*, 2000

[6] Scheidler C., Heiner G., Sasse R., Fuchs E., Kopetz H., Temple C.: “Time-Triggered Architecture (TTA)”, *EMMSEC'97*, Florence, Italy, 1997