Error Detection and Reconfiguration in Reliable Ethernet Train Networks

Hassanein H. Amer1, Magdi S. Moustafa2, Mai Ibrahim1, Ramez M. Daoud1

1Electronics Engineering Department, American University in Cairo, Cairo, Egypt
2Mathematics and Actuarial Science Department, American University in Cairo, Cairo, Egypt
E-mail: hamer@aucegypt.edu

Received July 26, 2011; revised August 28, 2011; accepted September 15, 2011

Abstract

In this paper, a novel reconfiguration technique is developed in the context of a fault-tolerant Networked Control System (NCS) in two train wagons. All sensors, controllers and actuators in both wagons are connected on top of a single Gigabit Ethernet network. The network also carries wired and wireless entertainment loads. A Markov model is used to prove that this reconfiguration technique reduces the effect of a failure in the error detection and switching mechanisms on the reliability of the control function. All calculations are based on closed-form solutions and verified using the SHARPE software package.

Keywords: Fault-Tolerance, Gigabit Ethernet, Markov Model, Train Control Network, Reliability, Coverage, Transportation Systems, Ethernet in Control

1. Introduction

One of the most popular wired network communication protocols is Ethernet1. Since its first release, it has been enhanced several times. Starting by the traditional Bus topology and the coaxial links, now Switched Ethernet architectures with optical links are available with speeds reaching 10 Gbps. Ethernet is based on the CSMA/CD mechanism. This is a non-deterministic protocol. With the introduction of switches, the non-deterministic nature of Ethernet is partially resolved. Now, different problems exist such as queuing delays and queue lengths. Even though Ethernet is non-deterministic by nature, this did not stop researchers in academia and industry from using the Ether-Channel as a communication medium for the most critical applications: Control Systems.

One of the most popular Networked Control Systems (NCS) protocols is CAN [1,2]. It was developed by BOSCH. Its function is to communicate control data from different control nodes in order to replace the traditional point-to-point links present in the early control systems. The automotive industry is the principal driving force behind the development of new control schemes. This is why CAN has a special version for automotive on-board network implementation. Since Ethernet appeared in the world of wired communication systems, the implementation of Ethernet as a communication medium for NCS was a must. The non-deterministic nature of Ethernet was first thought to be problematic because of the real-time constraints inherent in control systems; however, research showed that Ethernet (or IEEE Std 802.3) performed well in NCS either by changing packet format for real-time control messages, or by giving higher priority for these messages [3-5]. The standardization process for the use of Ethernet in control is also under way2. Rockwell Automation and the ODVA also proposed the EtherNet/IP as an industrial version of Ethernet and they have developed the CIP [6-9]. More references on this topic can be found in [10].

In [11], a new methodology was proposed, namely the use of Ethernet (IEEE 802.3) without any modifications in the context of NCS. This proved to be successful not only for pure control loads but also when mixing real-time and non-real-time messages. In [12], fault-tolerance was introduced on this scheme in the context of several machines working in-line.

This new methodology was also introduced in car networks [13]. It was shown that Gigabit Ethernet was able to integrate real-time control functions with non-real-time entertainment functions. More details about the use of NCS in cars can be found in [14]. It was also shown in [15,16] that the same principle is also applica-

---

1 IEEE 802.3 Standard.
2 IEC 61784-1,2 available at: www.iec.ch.
ble for train wagon control. In [17], two train wagons were studied; all sensors, controllers and actuators were connected on top of Gigabit Ethernet. It was shown that this architecture was successful in meeting the real time delay deadlines. Furthermore, the increase in system reliability due to fault-tolerance was calculated.

In this paper, the fault-tolerant network described in [17] is revisited. The effect of the efficiency of the error detection and reconfiguration mechanisms on the reliability of the control function is investigated. In order to reduce the effect of unsuccessful reconfiguration on system reliability, a novel scheme is developed. A Continuous Time Markov Chain (CTMC) is then used to prove that the reliability of this scheme is higher than that of a more conventional reconfiguration scheme.

The rest of this paper is organized as follows. Section 2 summarizes some of the work done in Ethernet train networks. Section 3 focuses on the new fault tolerant scheme developed in this paper and presents the Markov model that is used to calculate system reliability. In Section 4, it is proven that this new scheme increases system reliability. Finally, Section 5 concludes this research.

2. Ethernet Train Control Network

Due to the current technological advancement, entertainment and multimedia are becoming a necessity on board of moving vehicles [18]. Consequently, Ethernet evolves as a promising technology in train control networks over the currently used protocols such as Local Operating Networks (LonWorks), Train Communication Networks (TCN) and Controller Area Network (CAN) [19,20]. In [15], it was proven that the use of Ethernet, as a control protocol in trains, could allow carrying an entertainment load on top of the control load. This was achieved without jeopardizing the packet end-to-end delay requirement of the control data. A Gigabit Ethernet network model is proposed as a control and entertainment network within a one 60-seat train wagon [21]. The network consists of 250 nodes, the maximum number of sensors and actuators currently allowable in train standards [22]. Additionally, there are two categories of entertainment traffic added to the control traffic. The first load is in the form of video streams. The second load is a WiFi traffic produced from mobile wireless nodes (laptops).

With a packet payload of 32 bytes, the sensor to actuator packet end-to-end delay was measured using OPNET simulations. This measured delay includes all the processing, propagation, encapsulation and de-capsulation delays. This architecture succeeded in meeting required deadlines. All simulations were run for 16ms and 1ms sampling periods. More information can be found in [15,24].

2.1. Enhancing Train Network Reliability

In order to increase system reliability, two controllers are used instead of one controller [16]. A Control Server (Controller) handles the control load and an Entertainment Server handles the entertainment traffic (video streams and WiFi load). The Entertainment Server acts as a backup for the Controller in order to enhance system reliability. Figure 1 below shows the enhanced network model that was successfully simulated with OPNET.

2.2. Ethernet in Two Train Wagons

In [17], the network model is upgraded to include two wagons. The two wagon network consists of two Gigabit optical fibre star topologies, one per each wagon. In each wagon, the same network model proposed in [16] and shown in Figure 1 is modelled. These two star networks are connected to each other via a 10 Gigabit Ethernet optical fibre cable at the main switch level. Thus, the two wagons can exchange information.

To further increase network reliability, both Controllers and both Entertainment Servers serve as backups in case of a Controller failure. The worst case scenario is the one where three out of the four Controllers/Entertainment Servers fail; the remaining Controller/Entertainment Server handles the control load of both wagons while the entertainment is dropped in both wagons. Consequently, each sensor in the system has to multicast four replicas of its packets. These packets are sent to two Controllers and two Entertainment Servers.

In the context of two wagons, the main metric under study is the maximum sensor-to-actuator packet end-to-end delay. This measured delay includes all the processing, propagation, encapsulation and de-encapsulation delays. OPNET simulations showed that this architecture, when fault free, is successful in meeting required deadlines. The worst case scenario was also successfully simulated with OPNET; one controller handled the control load of both wagons while the entertainment was completely dropped.

3. Novel Error Detection/Reconfiguration Scheme

The fault-tolerance mentioned above is expected to increase system reliability. Let $R_{\text{control-FT}}$ be the reliability of the control function of both wagons. Furthermore, assume that the controllers in both wagons are identical. The same assumption will also be valid for the Entertainment Servers in both wagons. Let $R_K$ be the reliability of any of the two controllers ($K_1$ and $K_2$) and $R_E$ be the reliability of any of the two Entertainment Servers ($E_1$ and $E_2$). The reliability $R(t)$ is defined as the probability that a Controller/Entertainment Server is functional at time $t$. Hence

$$R_{\text{control-FT}} = 1 - (1 - R_K)^2 (1 - R_E)^2 \quad (1)$$

Note that $(1 - R_K)$ is the unreliability of a Controller while $(1 - R_E)$ is the unreliability of an Entertainment Server. $(1 - R(t))$ is the probability that the system has failed at time $t$. Intuitively, the fault tolerant architecture described in the previous section should increase the reliability of the control function in the context of two wagons. Without any fault tolerance, the control function will fail as soon as either of the two Controllers fails. Let this architecture have a reliability $R_{\text{cont}}$.

$$R_{\text{cont}} = R_K^2 \quad (2)$$

The time to failure of electronic equipment has been historically assumed to be exponentially distributed [24, 25]. The failure rate will therefore be constant. Let $\lambda_K$ be the failure rate of any of the two Controllers and let $\lambda_E$ be the failure rate of any of the two Entertainment Servers. The relation between the reliability and the failure rate is as follows [24,25]:

$$R_K(t) = e^{-\lambda_K t}; \quad R_E(t) = e^{-\lambda_E t} \quad (3)$$

By comparing $R_{\text{control-FT}}$ and $R_{\text{cont}}$, it was shown in [17] that fault-tolerance had increased system reliability as expected.

This increase in reliability relies on the implicit assumption that the switching tasks from a failed Controller/Entertainment Server to another operational Controller/Entertainment Server will always be successful. Next, the details of this switching mechanism are explained in the context of a $K$ failing and $E$ taking over its tasks. In the fault-free situation, all packets sent from the sensors are received by both servers: $K$ and $E$. Only $K$ responds to these packets, calculates the necessary control packet and sends it to the designated actuator node. A watchdog in the form of “live” packets is continuously exchanged between $K$ and $E$. When $E$ detects a missing watchdog (which indicates the failure of $K$), it gets into the loop to replace the inactive $K$ and sends the control packet to the designated actuator. The control procedure running on $E$ and used to backup $K$ in case of failure, must be designed to accommodate the loss of one packet. Also, the control system must not be susceptible to the loss of one control packet. This is to overcome the probability to lose, at most, one packet during the switchover between $K$ and $E$. A trivial solution in this case would be the “keep previous sample” technique. In this procedure, the actuator applies to the plant the previous action until a new control word is received.

This switching mechanism is susceptible to failure. For example, if the inter-communication between $K$ and $E$ fails, $E$ will assume that $K$ has failed and hence, will take over its tasks. Such a conflict will cause a system failure. More details about unsuccessful reconfiguration can be found in [26]. Furthermore, the reconfiguration process in control systems is covered in [27].

Since the success of the reconfiguration process is not guaranteed, it has to be taken into account in the reliability model. In the literature, the probability of successful detection/reconfiguration is called coverage [24,25,28]. The coverage is a parameter determined by the user and incorporated in reliability/availability models. It is known that a small mistake in the calculation of the coverage can lead to misleading reliability/availability estimations [25]. Also, system reliability is expected to decrease with a decrease in coverage.

A reconfiguration scheme is described next that aims at reducing the effect of coverage on the reliability of the control function $R_{\text{control-FT}}$.

Details of the New Scheme

Let $K_1$ and $K_2$ be the Controllers in the two wagons. Also, let $E_1$ and $E_2$ be the two Entertainment Servers. If one of the two controllers $K_1$ or $K_2$ fails, the other controller carries the control load for both wagons. Also, as a precautionary measure, the entertainment load is shut down as soon as one of the controllers fails. If one or both entertainment servers fail while both controllers are still operational, the entertainment is simply dropped without any need for reconfiguration. This strategy is
expected to produce a higher reliability when compared to a conventional strategy where the failure of any controller/entertainment server necessitates system reconfiguration. A Markov model is developed next to calculate system reliability based on the strategy described above.

Figure 2 shows this Markov model. The name of any state indicates the operational components in that state. Remember that $\lambda_E$ is the failure rate of the Entertainment Server and $\lambda_K$ is the failure rate of the Controller. The initial state is 2C2E. This is the error-free state; both controllers and both servers are operational. A failure of one of the controllers takes the system to state 1C2E. This transition will only occur if the reconfiguration is successful and the operational controller is able to take over the tasks of the failed controller and handle the entire control function in both wagons. This is why the transition rate from state 2C2E to state 1C2E is $2c\lambda_K$. If the reconfiguration is not successful, the control function fails and the system moves to state F (i.e., the control function failure state) at a rate of $2\lambda_K(1-c)$. Also, a failure of one of the Entertainment servers moves the model to state 2C1E. Since two servers can fail, the transition rate from 2C2E to 2C1E is $2\lambda_E$. The coverage does not affect this transition as mentioned above.

In state 1C2E, a failure of one of the two entertainment servers moves the system to state 1C1E at a rate of $2c\lambda_E$. The failure of the remaining operational controller takes the system to state 2E at a rate of $c\lambda_K$ and to state F at a rate of $(1-c)\lambda_K$.

In state 2C1E, the failure of C1 or C2 moves the system to state 1C1E; the transition rate is $2c\lambda_K$. If the reconfiguration is not successful, the system moves to state F at a rate of $2\lambda_K(1-c)$. The failure of the remaining Entertainment server takes the system to state 2C at a rate of $\lambda_E$. Here again, the reconfiguration process is not involved because the entertainment is shut off and the control function is not affected.

In state 2C, E1 and E2 have failed but the control function has not been affected since C1 and C2 are both operational. If either C1 or C2 fails, the model moves to state 1C; the control of both wagons is handled by the remaining operational controller. The coverage affects this transition and therefore, the transition rate from 2C to 1C is $2c\lambda_K$; also, the transition from 2C to F is $2(1-c)\lambda_K$.

Figure 2. Markov model.
In state 2E, both Controllers have already failed and one the Entertainment Servers is controlling both wagons. If either E1 or E2 fails, the control function is switched to the remaining operational server. Here again, the coverage is involved in the transition as shown in Figure 2.

In state 1C1E, the situation is more complex. One Entertainment Server has already failed as well as one of the Controllers. The remaining Controller is in charge of the control function of both wagons and the entertainment is turned off in both wagons. The remaining server acts as a hot stand-by for the remaining Controller. If the Entertainment Server fails, the control function is switched to the remaining server. This is why the initial state is called “4” and the one from 1C1E to F is (1-c)λK. Finally, in state 1E, the failure of the remaining entertainment server causes a system failure at a rate of λE. The same argument applies for state 1C where the failure of the remaining Controller causes a system failure at a rate of λK.

The system can be described by the Chapman-Kolmogorov equations. The row vectors \( dP/dt \) and \( P(t) \) are obvious and the transition matrix \( T \) is as shown above.

Given that the system starts in state 2C2E, these differential equations can be solved and the probabilities of being in each of the model states can be obtained in closed form. System reliability is the probability of not being in state F.

\[
P_{2\text{C}2\text{E}} = e^{-(\lambda_E + \lambda_K)}
\]

\[
P_{2\text{C}1\text{E}} = 2e^{-(\lambda_E + \lambda_K) t} \left(1 - e^{-\lambda_E t} \right)
\]

\[
P_{1\text{C}2\text{E}} = 2e^{-(\lambda_E + \lambda_K) t} \left(1 - e^{-\lambda_K t} \right)
\]

\[
P_{2\text{C}} = e^{-(\lambda_E + \lambda_K) t} - 2e^{-(\lambda_E + \lambda_K) t} e^{-2\lambda_K t}
\]

\[
P_{1\text{C}1\text{E}} = 4c e^{-(\lambda_E + \lambda_K) t} + e^{-(\lambda_E + \lambda_K) t} - e^{-(2\lambda_E + \lambda_K) t} - e^{-(2\lambda_E + \lambda_K) t}
\]

\[
P_{2\text{E}} = \left( c^2 e^{-2\lambda_K t} \right) \left[ 2(0.5e^{-2\lambda_K t} - e^{-2\lambda_K t}) + 1 \right]
\]

\[
P_{1\text{E}} = (4c) \left( e^{-(2\lambda_E + \lambda_K) t} \right) + (2c) \left( e^{-(\lambda_E + \lambda_K) t} \right)
\]

\[
- (2c) \left( e^{-(2\lambda_E + \lambda_K) t} \right) - (2c) \left( e^{-2\lambda_K t} \right)
\]

\[
- (4c) \left( e^{-(\lambda_E + \lambda_K) t} \right) + (2c) \left( e^{-2\lambda_K t} \right)
\]

\[
P_{1\text{E}} = \left( e^{-(2\lambda_E + \lambda_K) t} \right) \left[ X - \left( e^{-(2\lambda_E + \lambda_K) t} \right) Y \right]
\]

\[
+ \left( e^{-(\lambda_E + \lambda_K) t} \right) \left[ 2c^2 \right] - \left( e^{-(\lambda_E + \lambda_K) t} \right) \left[ 4c^2 \right]
\]

\[
- \left( e^{-2\lambda_K t} \right) \left[ c^2 \right] + \left( e^{-\lambda_K t} \right) \left[ W \right]
\]

where:

\[
X = \frac{4c^2 \lambda_K + 4c^3 \lambda_K}{2\lambda_K + \lambda_E}
\]

\[
Y = \frac{4c^2 \lambda_K + 2c^3 \lambda_E}{2\lambda_K + \lambda_E}
\]

\[
W = 2c^2 + 2c^3 + \left[ \frac{2c^2 (2\lambda_K + c\lambda_E)}{2\lambda_K + \lambda_E} \right] - \left[ \frac{4c^2 \lambda_K + c\lambda_E}{2\lambda_K + \lambda_E} \right]
\]

4. Efficiency of the New Scheme

In this section, it is proved that the novel reconfiguration scheme described above, increases system reliability (Rcontrol-FT). Conventionally, the entire system undergoes reconfiguration in the event of a controller/Entertainment Server failure. Such a system would be modeled by the CTMC depicted in Figure 3. In this model, it is assumed, for simplicity, that \( \lambda_K = \lambda_E = \lambda \). Consequently, the failure of any of the four controllers/servers may cause a system failure with a probability (1-c), where \( c \) is the coverage parameter discussed above. Any state in Figure 3 indicates the number of operational components in that state. A component can be a controller or an entertainment server. This is why the initial state is called “4” and the final (absorbing) state is called “0”. Moving from state “i” to state “i-1” (for i = 2 to 4) occurs when one of the operational components fails and the recovery is successful (with a probability c). Rcontrol-FT for this conven-
tional scheme is the probability of not being in state "0". To prove that the new fault-tolerant scheme presented in this research increases $R_{\text{control-FT}}$, both Markov models (the model for the new scheme in Figure 2 and the model for the conventional scheme in Figure 3) are solved and $R_{\text{control-FT}}$ is obtained. For simplicity, it is as-sumed that $\lambda_k = \lambda_e = \lambda = 1/\text{month}$. Figure 4 compares $R_{\text{control-FT}}$ for the two schemes. Two values of the coverage parameter are used: $c=0.95$ and $c=0.9$. For both val-ues, the reliability is higher for the novel scheme sug-gested in this research. Note that the difference between $R_{\text{control-FT}}(c=0.9)$ for the novel scheme model and $R_{\text{control-FT}}(c=0.95)$ for the conventional scheme, is very small. All calculations were verified using the SHARPE program4.

5. Conclusions

The use of Ethernet in Railway Networked Control Sys-tems at the sensor/actuator level is a relatively new research

area. Despite the fact that Ethernet is a non-determinis-tic protocol, it was proven that it would not violate re-quired real-time delays. This concept has been applied in industrial automation as well as in automotive environ-ments before its use in trains.

This paper focuses on the fault-tolerant aspect of a Networked Control System (NCS) in two train wagons. All sensors, controllers and actuators are connected on top of a single Gigabit Ethernet network. Furthermore, wired and wireless entertainment loads are carried on top of the same control network. Reliability is expected to increase because controller failures do not necessarily cause system failure. However, error detection and sys-tem reconfiguration need to be successful in order to improve reliability. The coverage parameter quantita-tively describes the probability of successful error detec-tion and reconfiguration.

A novel fault-tolerant scheme is developed. This scheme aims at increasing the reliability of the control function in the presence of the coverage parameter. A Markov model is then used to calculate system reliability. This reliability is then compared to that of a conventional fault-tolerant scheme with coverage. It is proven that the proposed scheme has a higher reliability. All results were compared to estimates produced by the SHARPE software package and were found to be identical.

6. References


