EMSOFT 2024 Home

wangshumo0102@gmail.com

#52 End-to-End Timing Analysis of Task Chains for TSN-based Distributed Real-time Systems

Main Edit Your submissions (All) Search

Email notification

Select to receive email on updates to reviews and comments.

▼ **PC conflicts**

Enrico Bini
Zhishan Guo

Rejected

Submission (618kB) Mar 31, 2024, 12:51:39 PM UTC · d8d52bd1

▼ **Abstract**

Distributed real-time systems are complex in nature due to their physical dispersion, typically composed of multiple electronic control units (ECUs). Such systems not only need to meet real-time constraints but also require restrictions on end-to-end timing to prevent adverse effects on the outcomes. We consider two main end-to-end timing semantics, i.e., maximum reaction time and maximum data age. The complexity of the system is compounded by the causal relationships often present between control tasks, where the output of one task leads to the input of another. The chains of such tasks also add to the complexity of end-to-end timing analysis.

The existing end-to-end timing analysis of task chains in distributed real-time systems typically employs the Controller Area Network (CAN) bus as the interconnect between ECUs. However, with the increasing volume of data, Time-Sensitive Networking (TSN) has emerged as a novel solution. In this paper, we investigate the end-to-end timing analysis of task chains in distributed systems based on the IEEE 802.1 QCR standard, constructs a model for task chain transmission over TSN networks, and analyzes the maximum reaction time and maximum data age. Combined with experiments, we show that our proposed method improves the performance.

▶ **Authors** (anonymous)

S. Wang, Y. Lin, Z. Feng, M. Shan, Q. Deng [details]

	OveMer
Review #52A	1
Review #52B	1
Review #52C	1
Review #52D	2

1 Comment: Administrator

You are an **author** of this submission.

Edit submission

Reviews and comments in plain text

Review #52A

Overall merit

1. Reject

Paper summary

The paper considers a distributed real-time system where ECUs are interconnected by a TSN network and flows are are scheduled, based on Asynchronous Traffic Shapers (ATS). A task chain is modelled by scheduling tasks executing on ECUs and communication tasks (one per hop) for the network part. The authors assume an event-triggered solution, I.e. a task is activated when the preceding one in the chain terminates. They also assume an implicit communication semantic: data are read at the beginning of tasks and written at the end. They propose an end-to-end timing analysis for the computation of maximum reaction time and data age. The analysis seems mainly based on previous work [13,27,28]. An evaluation shows some results, based on an automotive application.

Strengths

Interesting problem.

Weaknesses

Poor presentation, making the paper unclear.

Detailed comments

Timing analysis of task chains in a distributed systems based on TSN without synchronization makes sense. Indeed TSN is a very promising solution for real-time communication and synchronization induces complexity. Unfortunately the paper is poorly presented. There are many missing words and strange sentences, e.g. « A task chain C ={...} are satisfied ». Task and network models are presented with a huge amount of notations which make it difficult to follow. Task chain definition is a bit confusing. Event c0 is not in the description of the chain (C = {z, c1, ...}). The definition of data age uses cn+1 which does not exist. Case 1 in lemma 2 is very unclear. It is based on previous work which are not detailed. What is the author contribution on this part. Case 2 is also unclear. The rest of the analysis is more or less based on these two cases. In the evaluation part, formula 1 is mentioned. What is it? Obviously not equation 1 which concerns reaction time. Therefore I cannot decide whether the proposed solution is valid or not.

Review #52B

Overall merit

1. Reject

Paper summary

This paper studies the end-to-end latency analysis for TSN-based distributed real-time systems. The claim in this paper is that existing analyses used CAN bus for communication, which cannot be used when the communication fabric is TSN. This is a very strong statement and the paper does not have a support to explain why they cannot be modified or extended. The paper has a decent introduction of the end-to-end latency. But the technical part of the paper of Section V is not presented with correct scientific writing and the statements made in this paper are questionable, as no clear assumptions are made.

Strengths

- A timely topic.

Weaknesses

- The first 6 pages of the paper have very basic introduction of cause-effect chains and TSN. As there have been many recent papers about cause-effect chains and TSN, it should be sufficient to introduce the concepts needed to understand this paper and refer to existing results for very basic concepts. For example, the examples used for Figures 1-5 are overly basic and not really highly relevant to the technical contribution of this paper. I suggest a careful revision of the first 6 pages to highlight only the necessary part relevant to this paper.

- The analysis in Page 8 suddenly uses arrival curves, service curves, convolution operators, and deconvolution operators. This is VERY unclear. It is incorrect to put assumptions only in the proofs without proper justifications of these assumptions. I don't understand how the TSN model studied in this paper is translated into arrival curves and service curves, as there is no presentation about that in the paper. Therefore, the correctness of this paper is questionable.

- The data age analysis is only explained without any proof. The recent result in [12] shows that the data age and the maximum reaction time are equivalent in general. In case the behaviour of the system complies with the assumptions made in [12] for equivalence, the maximum reaction time is the same as the maximum data age.

Detailed comments

Some further minor comments:

- Direct communication is typically referred to as explicit communication in the literature.
- I understand the paper utilizing time-triggered data propagation. The following paragraph in Section V.B reads very strange to me, "*However, in the task chain model of this paper, since an event-triggered approach is used with consideration for fixed-size buffers and the possibility of data frame overflow, the time of the incentivized action is the release time of the next task in the data transmission of the task chain, which is also the end time of the last task in the task chain.*"

Review #52C

Overall merit

1. Reject

Paper summary

The paper analyzes the end-to-end latency (reaction latency and data age) of task chains spanning multiple ECUs. Communication between ECUs is realized by TSN, specifically with Asynchronous Traffic Shaping (ATS). In the used model, the first task in the chain is periodically activated, and later tasks in the chain are event triggered. All tasks are non-preemptive and scheduled on a single core under fixed-priority scheduling. The analysis is based on existing works that are then combined for the proposed analysis. Evaluation is performed on synthetic task sets.

Strengths

- By considering TSN-based distributed systems, the paper addresses a relevant area
- The provided background information helps to understand the different parts of the paper

Weaknesses

- The selected system model simplifies the problem of determining reaction latency and data age
- The novelty of the approach is limited as a large part of the analysis is based on related work
- The experiments are not detailed enough to demonstrate the performance improvement of the approach

Detailed comments

The background section of the paper gives a good introduction to ATS.

The abstract states that real-time constraints and end-to-end timing constraints must be met. I suggest writing that task deadlines, as well as end-to-end constraints, must be met. It is formulated in this way later on in the introduction.

The introduction describes initially how task chains are structured and how communication between tasks is realized. This part is not easy to follow, and I suggest to rewrite it for clarity. In particular, the part that describes register communication is very hard to follow. It could also be good to add a reference, for example [19].

When discussing previous end-to-end timing analysis that considered CAN, the references point only to works that focus on CAN analysis [7][8], not analysis of end-to-end delays on distributed systems with CAN.

All section numbers in the contribution list have an offset of -2.

The related work section seems unstructured. Several works that address similar task models are not included in the discussion. For example, [a] or [b].

[a] K. Tindell and J. Clark, ''Holistic schedulability analysis for distributed hard real-time systems,'' Microprocess. Microprogramm., vol. 40, nos. 2–3, pp. 117–134, 1994. [b] J. J. G. Garcia, J. C. P. Gutierrez, and M. G. Harbour, ''Schedulability analysis of distributed hard real-time systems with multiple-event synchronization,'' in Proc. 12th Euromicro Conf. Real-Time Syst. Euromicro RTS, 2000, pp. 15–24.

The task model starts by discussing two main communication forms, implicit communication and LET. Just afterwards, in the text as well as in Fig. 4, explicit communication is also discussed. Thus, it could be better to list all three directly. Explicit communication is also supported by AUTOSAR.

To better understand the task model it would be good to first explain that the terminology of scheduling tasks and network tasks is used to refer to the different parts of the task chain.

To me, it was unclear what exactly the input buffer of the task is or how it affects the execution of a chain.

The network model describes N ECUs, N flows and N token bucket shaping queues. Does this limit the applicability of the model, as it seems the number of possible switches between the nodes is limited by this.

The model further describes that the queue only selects the highest priority data frame for transmission without considering other possible interfering scenarios. This should be discussed more. Which scenarios are omitted?

In several places, "reaction time" is mixed up with "response time". For example, in the first sentence after the bullet list of definition 1.

Definitions 2 and 3 describe reaction time and data age, respectively. The definition is based on the release and finish time of jobs. However, the definition uses tasks instead of jobs in the formula. Instead, it should be described on the data path of communicating task instances, as for example in [12]. The data age is further described to finish at $r(c_{n+1})$, where c_n is the last event/task of the chain. Thus, it is not clear what the event after the last event in the chain is.

The stated execution time of the different tasks in the illustrative example is not correct. In the text, it is 3, but in the figure and based on the other values later, it should be 4. It also seems that Fig. 5 is not to scale. For example, the release of the second instance of task 1 is at t=6, but the dashed line is quite close to t=8. To me, it was also not clear why only every second instance of task 1 triggers its successor task. Based on the system model, there should be no rate transition in the task model. With this, the maximum reaction time should also be larger than shown. Assume the event arrives just after t=6. The instance at t=12 does not lead to a valid data path that reaches the end of the chain. Thus, only the next instance at t=18 picks the event and reaches the end of the chain with a max reaction delay of 42.

The steps in Equation 6 are not clear to me. Why are $t(c_1) - t(c_0)$ canceling out? It also seems the equation could directly be written as $t(c_0) - t(z)$. Since $t(c_0)$ is the end time of the event and, with that, the release of the first task.

Several parts of the analysis are well known from the literature. For example, the upper bound of the sampling interval (Lemma 1), e.g. [19].

The main part of the analysis was not clear to me. It seems big parts are from related work, i.e. [13] for tasks and [27,28] for messages. It would be very helpful to highlight better which parts are new contributions. The calculation of task response times (Case 1) seemed also to be based on the buffer size. To me, it was not clear how the buffer size is affecting the response times in the described model. Similarly, the bound on the network delay is described very briefly, and it is difficult to assess its correctness.

It was not surprising that, under the task chain model, data age and reaction latency are identical with the exception of the initial sampling delay in the reaction latency. Since no over- or under-sampling can occur, only one job chain exists from the sampling task to the last task of the chain.

The evaluation presents the results of the proposed analysis with varying system parameters. The data age and reaction latency results look very similar, even though the text states that this is mainly an effect of the normalisation. But looking at theorem 1 and 2, it would be expected that they are fairly similar as they only differ in the initial sampling delay for the reaction latency.

A comparison is provided against the Davare et al. bound where the worst case response time is upper bounded by the period. A comparison for example with the approach in [a] could be interesting (available in the MAST tool).

Minor Comments

- Abstract: "...constructs a model..." -> constuct
- Section 2: "...Durr et al in [10]define..." space after [10] missing
- Section 2: "Gunzel et al. in [11] study extended the definition..."
- Section 2: Sentence starting with "The event-triggered mechanism for task chains..." is missing parts
- Section 2: "In the work[5]a method" space missing after [5]
- Section 2: "Bahar et al. ..." Bahar is the first name of the author, also the citation to their paper is missing, i.e. [4]
- Section 3: Reference to IEEE 802.1Qcr is the bibkey in brackets
- Definition 6 refers to a definition IV-C.

Review #52D

Overall merit

2. Weak reject

Paper summary

The paper proposes a timing analysis of chained tasks that communicate using the ATS defined in the Time-Sensitive Networking standards and presents an evaluation of the analysis.

Strengths

The topic is very interesting and relevant to the EMSOFT audience. The motivation behind the timing analysis of task chains in Time-Sensitive Networking systems is clear.

Weaknesses

In general, the paper lacks of readability. Most of the background approaches are not clearly explained and some assumptions in the analysis are not justified. The evaluation lacks sufficient detail to be convincing.

Detailed comments

In general, the paper is not easy to read. For example:

- from the text in Sect. III it is not easy to understand how the ATS works. The description and the explanation provided by Algorithm 1 is is too essential and lacks clarity.
- Formula (6) is not clear, for example, it is unclear what "– r(c1) + r(c1)" in the third row means.

Moreover, there are some assumptions in the network model that are not convincing and deserve more insights.

In particular, the following one: "When a data frame passes through the shaping queue and qualifies for transmission time, the transmission selection algorithm will poll the head of the queue to select the data frame for transmission. In our research, we consider that at this point, only the data frame with the highest priority in the queue is selected for the next transmission step, without considering other possible interference situations". The Authors should explain why other possible interference situations are not considered and to what extent this assumption generally holds.

The analysis is difficult to follow, mainly due to the high number of symbols. A Table with the symbol description would be beneficial to the reader.

The evaluation results are not fully convincing, for the following reasons.

1. What is the evaluated network topology? The network delays are heavily affected by topology.
2. It is not clear if the evaluation was made on an experimental test-bed or using simulations.
3. The normalized results, i.e., Normalized Data age bound and Reaction Time Bound are not defined. The y-axis in Fig. 7d does not show the units.

Minor issues

- The language should be revised. For instance: -- the following sentence of the abstract should be fixed "In this paper, we investigate the end-to-end timing analysis of task chains in distributed systems based on the IEEE 802.1 QCR standard, constructs a model for task chain transmission over TSN networks, and analyzes the maximum reaction time and maximum data age." Moreover, the standard is an amendment of the IEEE 802.1Q, not a standalone standard, so "cr" should be used instead of "CR". -- the paragraph below Fig. 1. "Currently, there have been studies combining TSN with task chains. [9] Modeling task chains using real-world automotive cases and using the IEEE 802.1Qbv protocol as a bridge for network transmission." -- and others.

- Some statements should be supported by references or explanations. For instance, in Sect. I "during automatic vehicle cruising, if the reaction time of the control unit exceeds 50 milliseconds, although it may still complete the deceleration control within the deadline, the delay in the control signal could result in a sudden deceleration and loss of vehicle stability". Where does the 50ms reaction time comes from? Does it depend on the vehicle speed?

- In Sect. I. the Authors should explain what "an incentivized output" is.

- Typo in the symbol above Formula (1).

@A1 Administrator · 5d

Unanimity on a *rejection* decision has been reached without a discussion during the TPC meeting.
