Higher busloads for automotive CAN network clusters

Nicolas Navet and Hervé Perrault

Authors
Nicolas Navet
Inria/RTaW
615, rue du
Jardin Botanique
FR-54600
Villers-les-Nancy

Hervé Perrault
PSA Peugeot-Citroën
18, rue des Fauvelles
FR-92256
La Garenne-Colombes
Cedex

Link
www.loria.fr/~nnavet

Sometimes, a 60 % loaded CAN network can be more efficient than two 40 % loaded CAN segments interconnected by a gateway causing delays and jitters. The first obvious way to optimizing a CAN system is to keep the amount of data transmitted to a minimum, specifically limits the transmission frequency of the frames. This requires a rigorous identification and traceability of the temporal constraints. Given a set of signals or frames, and their associated temporal constraints (freshness, jitters, etc.), there are in addition a few configuration levers than can be triggered:

- Desynchronize the stream of frames by using offsets (see Figure 1).
- Reassign the priorities of the frames, so that the priority order better reflects the timing constraints.
- Re-consider the frame-packing (i.e. allocation of the signals to the frames and choice of the frame periods, so as to minimize the bandwidth usage while meeting timing constraints).
- Optimize the ECU communication stacks so as to remove all implementation choices that cause a departure from the ideal CAN behavior.

Configuration and verification algorithms used for the first three items have to guarantee the temporal behavior of the communication system, and ideally be optimal, or provide lower-bounds on their efficiency. In our view, a busload threshold for an "easy" CAN cluster integration is around 35 % to 40 %, and below this limit, the latencies and freshness constraints are rather easily to "manage". Overcoming this limit implies more detailed supplier specifications on the one hand, and, on the other hand spending more time and effort in the integration/validation phase.

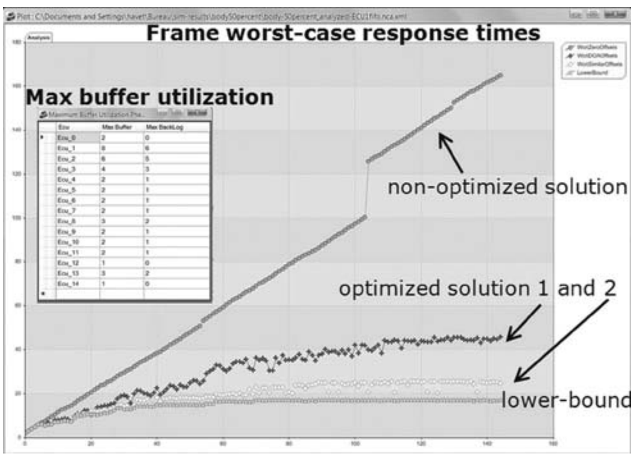

Figure 1: Screenshot of a Netcar-Analyzer showing maximum buffer-utilization and CAN frame worst-case response times (by decreasing priority) for different offset configurations. This graph shows the typical gain one can expect with offsets.

Simulation versus analysis

Early in the development cycle, when ECUs are not available, simulation models and analytical models are the two possible verification techniques. Both provide complementary results and, most often, none of them alone is sufficient. On the one hand, numerous experiments suggest that simulation alone is not appropriate to find the worst-case scenarios because they are too rare (see Figure 2). On the other hand, worst-case analysis cannot help to quantify how rare these events are, nor how long they last, nor what the average (or any other relevant statistics) of the response times are.

However, it is possible to derive by analysis the phasing conditions between ECUs, specific to each frame, that cause its worst case response time. Then, using a simulation tool, it becomes possible to observe for how long this situation lasts and where the ECU clock drifts lead from there. Such simulations also contribute to validate the results obtained from the analysis tool (see Figure 3), which is needed because these tools are usually commercial black boxes and, though progresses are steadily being ▷

made, they have to make simplifications about the hardware and the communication stack. Besides, because of the complexity of the schedulability analyses, there is always the risk that the tool implementation or even the analysis itself is flawed, as it turned out to be the case with the basic CAN schedulability analysis.

There are now COTS (commercial of-the-shelf) tools to support the verification activity, even freely available tools such as RTaW-Sim for simulation and Netcar-Analyzer for schedulability analysis. For CAN, analysis consists mainly of schedulability analyses, providing upper bounds on the considered performance metrics: latencies, transmission jitters, size of the waiting queues at the ECUs and gateway levels, etc.

Optimized CAN networks means higher busloads, and indeed they may now easily exceed 50 % of load. But because there is less slack, there is a need for models that are more fine-grained than they were in the past. In particular, models should account for transmission errors and possibly ECU reboots. Additionally, they should consider the use of a periodic communication task responsible for building the frame and issuing the transmission requests. In some cases, this frame may suffer delays caused by higher priority activities. Possible asynchronisms between the applicative level tasks that produce the signals and the communication task needs to be evaluated. Sometimes such delays can be larger than the latencies on the CAN network. More fine-grained models of the hardware and communication stack are necessary. For instance, taking into account the ECU clock drifts may change drastically the conclusions that can be drawn from a simulation. The same holds for a worst- ▷

Introduction
When CAN was introduced, the busloads were limited and the specifications of the communication stack features, priorities and periods, etc. were defined more to handle scalability and overcome microcontroller limitations than bandwidth optimization. Optimizing CAN networks, which includes reaching higher load levels, has now become a requirement for several reasons. It helps to master the complexity of network architectures, reduces the hardware costs (weight, space, consumption, etc.), and facilitates an incremental design process. Additionally, it may avoid the effort, the risk, and the time to master new technologies.

Conclusion

"We can consider that when an application requires more than three or four CAN networks, it could be a better choice to introduce a new networking technology. As the most important needs for CAN bandwidth come from powertrain and chassis domains, a 'natural' technology could be Flexray. Another communication technology, which should be considered to increase the bandwidth is the recently introduced CAN-FD from Bosch. It may provide a good trade-off between the difficulty of the migration path and additional bandwidth availability. Nevertheless, in many cases, optimizing the normal CAN networks will help to defer the introduction of new technologies, at least for a subset of car domains. Using CAN at higher load levels requires however additional time and effort, be it for the supplier specifications or the verification. But in our view the current state of the technical literature on CAN and the COTS software tools are now mature enough to alleviate this additional work and succeed in building truly safe optimized CAN-based communication systems."

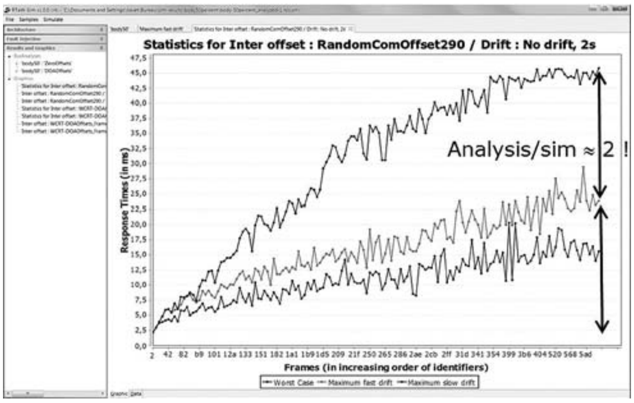

Figure 2: Worst-case response times (by decreasing priority of the frames) obtained by analysis (black curve) versus maximum values collected during long simulation runs for two ECU clock drift values (screenshot of RTaW-Sim).

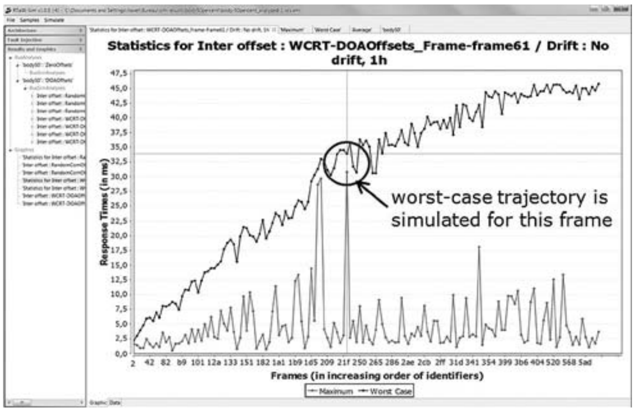

Figure 3: Worst-case response times (by decreasing priority of the frames) obtained by analysis (blue curve) versus maximum values collected by simulation. The trajectory that was simulated here is the one leading to the worst-case response time for a specific frame. As the black circle shows, the worst-case response time for that frame is close to what can be obtained by simulation.

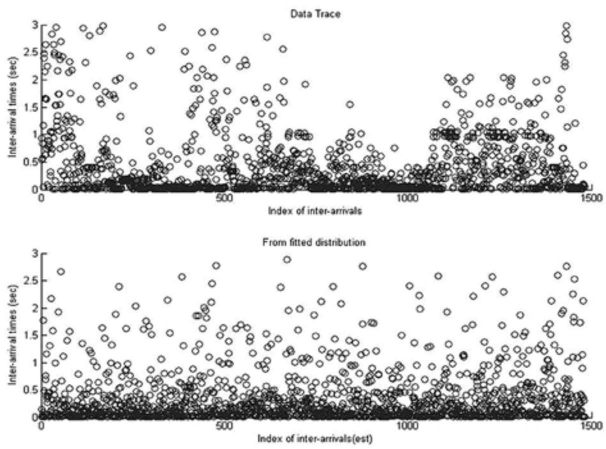

Figure 4: On the two graphs, the X-axis shows the index of the aperiodic frames while the Y-axis is the time between two successive aperiodic transmissions. The upper graph is a real data trace collected while driving (only the aperiodic frames). The lower graph is an artificial data trace generated with a probabilistic model of the aperiodic frames (here Weibull interarrivals with parameters fitted with maximum-likelihood estimation using the real data trace). The probabilistic model can be used both for simulation and worst-case analysis.

case schedulability analysis, when explicitly modeling a FIFO waiting queue.

Characterization of the traffic is another topic, especially the non-periodic part of the traffic and the transmission jitters. The non-periodic traffic is generally difficult to characterize, but if overlooked, one will tend to underestimate the frame latencies, which may have an impact on the safety.

Departure from the ideal CAN behavior

Up to rather recently analytical models were often much simplified abstraction of reality: Usually overly pessimistic (e.g. regarding the non-periodic traffic) and sometimes even optimistic, which means unsafe in our context. Indeed not all the classical assumptions made on the ideal CAN scheduling model are met by the implementations. Examples include:

- Non-abortable transmit request (some communication stacks/controllers may not offer the possibility to cancel lower-priority transmission requests, when a higher priority frame is released),
- Limited number of transmit buffers,
- Delays in refilling the transmit buffers,
- The use of a FIFO waiting queue for frames, or any other policy than the Highest Priority First (see Figure 5),
- Internal CAN controller message arbitration based on transmit buffer number rather than CAN-ID,
- Frame queuing not done in priority order (but for example by PDU index in Autosar) because of the communication stack.

Whether or not the CAN communication stacks will depart from the ideal CAN behavior may make in practice a large differ- ▷

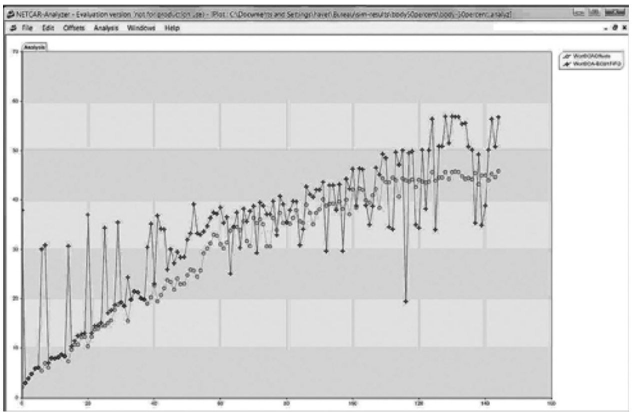

Figure 5: Frame worst-case response times by decreasing priority on a typical body network. The blue curve shows the results when all nodes have prioritized waiting queues for the frames. The blue curve shows the actual worst-case response times when there is one station (out of 15) that possesses a FIFO waiting queue. As one can observe, in the latter case many high priority frames will suffer more delays, and potentially they may not respect their timing constraints (e.g. deadline, jitter in reception).

ence in terms of performance and predictability. For instance, a single station with a FIFO queue can create bursts of high priority frames that will impact the latencies of the frames sent by all the other stations (see Figure 5), possibly it may even propagate to other networks through the increased jitters of the frames that are forwarded through the gateways. In a general manner, if the system designer does not have control over the communication stacks of all the ECUs that make up a system, he should use conservative assumptions for the validation. Fortunately, since a few years and the identification of a flaw in the original CAN schedulability analysis, significant progresses have been made in our view and the main issues have been identified and accounted for in the schedulability analysis.

Better adherence to the CAN priority behavior, can be enforced by more detailed and more constraining specifications for the suppliers. Also, to some extent, tools such as the RTaW-TraceInspector can perform the verification by means of analyzing transmission traces. ◀

More information

This excerpt derives from the iCC paper by the same authors ("CAN in Automotive Applications: a Look Forward") available on CiA's website (www.can-cia.org) in PDF format.

Related articles

"CAN with flexible data-rate" by Florian Hartwich, page 10 and following in this issue.