IEC 61508 Case Study

Peter Bernard Ladkin

RVS White Paper 3
20 February 2013

This note proposes a series of exercises concerning the assurance of safety-related systems involving E/E/PE components using the principles of IEC 61508. The exercises have been commented by others in a mailing-list format. References to those comments in the mailing list archive are given. I don't consider that any of the questions posed have yet been answered adequately.

The Situation

Suppose my customer has a large area of ground for critical infrastructure, say a square kilometer, which she wishes to maintain under constant surveillance with video and infrared sensing. The spatial geometry of the area with its structures is complex enough that fixed-base surveillance is impractical; also she judges that potential attackers could too easily survey and selectively disable ground-based sensing. She has decided for aerial surveillance and obtained the necessary airspace restriction, airspace over her domain, extending a further half kilometer beyond the boundary in each direction, in which flight is prohibited, except for her surveillance devices, under 2,500 ft altitude. This infrastructure lies in a built-up area (it didn't when it was originally built, but the city has grown up around it) and includes containers holding dangerous bioactive substances that must not be released into the atmosphere.

She has bought aircraft which were originally personal transports but have been modified as unmanned aerial vehicles (UAVs) with a variant of the proven-in-use automated flight systems (AFS) of the Boeing 737, the world's most highly-flown jet airliner, with of course the flight parameters adjusted for the rather smaller, slower surveillance vehicle.

The autothrottle of this airliner has a mode called RETARD, into which it passes when the aircraft is to flare on landing. In RETARD mode, the autothrottle closes the throttles to idle. Its sensors are radio altimeters (also known as radar altimeters), RAs. The Boeing 737 slaves the autothrottle to just one RA, but since RAs are not sufficiently reliable, three are included in the drone and a 2oo3 voting module has been added to the autothrottle software.

This modification is similar to what some safety experts suggested after the Turkish Airlines accident at Amsterdam Schiphol in February 2009 [ONZ 2009, ONZ 2010].

Some brilliant systems engineers have produced a safety case for the infrastructure in which the acceptable estimates of dangerous failure of the infrastructure are predicated on the surveillance function being on hand with a failure probability of 1 in 10^(-9) op hours, because they had heard that that is the touchstone for catastrophic single-cause failure of commercial aircraft systems and the drone will be using (at least) one, namely the AFS.

Furthermore, because the drones have been included as part of the overall system, as well as airworthiness requirements it has to fulfil IEC 61508 requirements, because the overall plant contains E/E/PE subsystems, such as the AFS of the drones, and thus lies under that remit. They must also fulfil the aerospace requirement as well, for example, because a drone takes off and lands at an airport outside the restricted airspace that "belongs" to the plant.

Say I have the contract to build the 2oo3 voting software. I'm going to do it with SPARK and let's imagine that I'm very good at it and known to be so. This is a modification to an airworthiness-approved AFS and I've done it using formal methods so let's imagine the CAA is happy.

Airline pilots see RA failures as unexceptional events, indeed "common enough" as one has said to me. A pilot flying 80 hours a month over a 35-year career gets about 35,000 hours flight time. Let us suppose that about one out of three has an RA fail on her, so no matter what the manufacturer says let's peg the reliability at one failure in 100,000 hours, or 1 in 10^(-5) per hour. We have three. It follows that chances of independent failure of at least one (that is, not couting common-cause failures) are 3 x 10^(-5) per hour. When an RA fails, it usually outputs 0, so in fact even if my SW is perfect there is a 1 in 10^(-10) chance per hour that my voting algorithm will (correctly for it) see two saying ground is near and tell the AFS so, which will shut off the throttles and - independent of the fact that the airplane will drift down into a built up area and likely cause harm - cause my client's surveillance function to fail.

The AFS supplier has to deal with that 10^(-10) issue, not me. I don't worry about it (turns out he doesn't have to either, when the airplane is on station - see below). I have also proved formally (because I used SPARK) that when all three RAs agree, my SW correctly passes that result to the AFS. The AFS supplier also has to worry about the 1 in 10^(-15) case that all three break (again excluding common cause), but he won't really have to do that, it turns out.

Then there are the other cases. The dangerous failure case I have left to worry about is that my voting algorithm will see two correct RAs saying 2,000 ft, one failed RA saying 0 ft, and decide 0 ft. Now, my client has presented a safety case to the regulator predicated on a surveillance failure event probability of 1 in 10^(-9) but that's not my problem. What may be my problem is that my drone drifts down with substantial gas in the tank into the neighborhoods surrounding the plant, hits something, explodes into flames and hurts someone. That counts as a dangerous failure of the plant, since the surveillance function is integral to the plant. My client, the plant owner, is planning the drones to run with essentially the same systems for 20 years, so at 10,000 hours in a year we are looking at 200,000 hours. She conducted a survey in the neighborhood, which thinks that even one accident is too much. So she suggests to them: how about we say it will be more likely than not that no accident will occur? The neighborhood says OK. Acceptable risk is thus 2.5 x 10^(-6) pdfh.

Not all possibilities of drift-down accidents can be attributable to my voting SW making a mistake, but if it does, it will be the false 1oo3 decision going in favor of the failed device rather than the other two.

My client has to show the regulator that the chances of a drift-down occurring are less than 2.5 x 10^(-6) ph. The regulator says: OK, enumerate the ways in which that can occur (causes). Simple arithmetic says that each individual cause can occur with a frequency at most 2.5 x 10^(-6) ph. Show it.

She asks me for an answer, because my software is one potential single cause of a drift-down. I can answer:

(a) the 2oo3 voting software was developed by experienced, reliable people using state-of-the-art SPARK methodology.

(b) The 2oo3 voting software gives the correct decision more than 96% of the time.

We can imagine the regulator will be happy with answer (b). But I give answer (a), which she passes on to the regulator. The regulator says, sorry, I need to know inter alia that the chances of a drift-down occurring are less than 2.5 x 10^(-6) ph, and you haven't shown me that. I cannot approve your system according to IEC 61508 until you do.

She takes the regulator to court, on the basis she has given a satisfactory safety case according to the applicable standard IEC 61508 and her system should be approved.

Questions 1:
You are the judge.
Do you decide for my client, or do you decide for the regulator? On what basis?

Further Development

My client has not started up her plant since she has not yet obtained approval to operate and is concerned about potential criminal liability if an accident happens and she is found not to be conformant with IEC 61508, and so is her insurance company. And the legal proceedings are costing her money. Consequently I have not been paid for the work that we did on the 2oo3 voting SW.

The court has ordered her, the plaintiff, and the defendant, and the regulator to try to reach an agreement. My client has asked me for the documentation that the code was developed according to state-of-the-art principles using state-of-the-art critical-software development environments, and the appropriate amount of great care.

Now, I am rather annoyed at not being paid. I point out to her that providing the appropriate documentation will cost us resources for which I am in some doubt we will be recompensed, given that she hasn't paid for the completed SW. I present her with our ISO 9000 certification (which costs me nothing) and suggest she assume we developed the software by magic.

My client goes to the regulator and says "Look, even by your argument, were it to be valid (which I do not admit), this voter only has to be correct 99 times out of a hundred times it is asked to judge a 2-against-1 discrepancy. By your own argument, that is a requirement above even that of SIL 1. And SIL 1 is the highest dangerous-failure frequency range at which **any** justification is required by 61508 that the safety function achieves it. So even if SW gets target failure frequencies as you claim, but which I do not admit, there is no requirement to show that this SW achieves the target failure frequency you want to see it achieve. So you are obliged to admit that my system is IEC 61508-conformant."

The regulator says: "You must be joking! You want me to bless a system, without giving me any information about how this safety function behaves? I can't do that. I have a duty of care."

They end up back in court. My client argues that, even on the regulator's interpretation that a target probability of dangerous failure per hour is appropriate for a pure software safety function, the regulator already has all the documentation required for him, under his own interpretation, to declare that the system is IEC 61508-conformant.

The regulator argues that this is absurd. He has a duty of care to the public and he hasn't seen anything that tells him with any degree of confidence how this software even behaves.

My client argues that the duty of care is a different matter. This case is about whether the regulator is to declare that her system conforms to IEC 61508 and she is claiming that he must do so, even by his own argument (which, she reminds us, she disputes).

Questions 2:
You are the judges. What do you decide? On what basis?

Even Further Development

It turns out that my client isn't as stubborn as I had surmised, and is fairly certain she will reach accommodation with the regulator some way or another. She feels my pain at not being paid, so has her solicitors draw up a legal charge for my agreed fee plus 40%, payable one year hence or on the first day the plant is in service, whichever comes first. She explains that the 40% is partly for the delay, which of course has nothing to do with me, and partly for the documentation that our SW does in fact make the correct decision in 1-against-2 cases 98% of the time (that's a margin of two over what the regulator agrees is needed). We sign.

The plant goes into service. In reviewing the documentation we produced in order to make a bid to another client, I notice that the 96%-correctness figure which was mooted, and indeed the 98% figure which we justified, does not suffice for the purpose for which it was intended. 99% would suffice. I start work on the tightened reliability case.

Question 3:
Why is neither 96% or 98% sufficient, but rather 99%?

I inform my client, who informs the regulator. But 30 minutes afterwards, before anyone has any time to do anything, there is a drift-down incident and the aircraft lands on somebody's house, but still inside the restricted airspace, with serious consequences.

The inquiry concludes conclusively that it was due to a faulty RA and my SW deciding, wrongly, for the faulty RA in a 1-against-2 situation.

The regulator decides to prosecute my client, on the grounds that the system was not developed according to IEC 61508. Yes, we had given a safety case, but that safety case was faulty. It had proposed and shown an inadequate criterion for component reliability of part of a safety function. My client, says the regulator, has been negligent.

My client argues that she had provided the regulator with all the argument and documentation he wanted, in good faith, even though she contends that that was in fact more than what the IEC 61508 standard requires. There was a mistake in the safety case, in a particular reliability assessment that she claims was not required for adjudication of IEC 61508 conformance, but which she provided to the regulator in any case, because they had asked for it and she had it. Her system was indeed IEC 61508 conformant, contrary to what the regulator claims, and she rejects the claim that she has been negligent.

The regulator replies that he agrees that the requirement is that she is to have shown IEC 61508 conformance, but contends she had not done that.

She says to the regulator “but you thought, even according to your own criteria, that I had!” The regulator replies that that is correct, but she misled him, unintentionally, and he mistakenly agreed that the system was conformant. In fact it wasn't, because he had been misled. And that is what he is contending. There is legal precedent, all agree, that, if one has not conformed with IEC 61508 and an accident happens, the developer has been negligent (although negligent to what degree is determined on a case-by-case basis).

Questions 4:
You are the judge. What do you decide? On what basis?

Acknowledgements

This note derives from two submissions of mine to the University of York Safety-Critical Systems Mailing List (as it then was) on 21 and 22 May, 2009 [Ladkin 2009.1, Ladkin 2009.2]. I am grateful for replies from Alan Bain [Bain 2009], Thierry Coq [Coq 2009], Andy Loebl [Loebl 2009.1, 2009.2], Bertrand Ricque [Ricque 2009.1, 2009.2, 2009.3, 2009.4] and Rolf Spiker [Spiker 2009].

References

[Bain 2009] Alan Bain, Note to York Safety-Critical Systems Mailing List, 29 May 2009. Available at http://www.cs.york.ac.uk/hise/safety-critical-archive/2009/0172.html

[Coq 2009] Thierry Coq, Note to York Safety-Critical Systems Mailing List, 26 May 2009. Available at http://www.cs.york.ac.uk/hise/safety-critical-archive/2009/0169.html

[Ladkin 2009.1] Peter Bernard Ladkin, Note to York Safety-Critical Systems Mailing List, 21 May 2009. Available at http://www.cs.york.ac.uk/hise/safety-critical-archive/2009/0160.html

[Ladkin 2009.2] Peter Bernard Ladkin, Note to York Safety-Critical Systems Mailing List, 22 May 2009. Available at http://www.cs.york.ac.uk/hise/safety-critical-archive/2009/0163.html

[Loebl2009.1] Andy Loebl, Note to York Safety-Critical Systems Mailing List, 22 May 2009. Available at http://www.cs.york.ac.uk/hise/safety-critical-archive/2009/0165.html

[Loebl 2009.2] Andy Loebl, Note to York Safety-Critical Systems Mailing List, 26 May 2009. Available at http://www.cs.york.ac.uk/hise/safety-critical-archive/2009/0167.html

[ONZ 2009] Dutch Safety Board (Onderzoeksraad): Preliminary Report (Accident, Accident Boeing 737-800, TC-JGE, 25 February 2009, Amsterdam Schiphol airport), 2009. Available at http://www.onderzoeksraad.nl/docs/rapporten/Prelimenary_engels.pdf

[ONZ 2010] Dutch Safety Board (Onderzoeksraad): Final Report, Crashed during approach, Boeing 737-800, near Amsterdam Schiphol airport, 25 February 2009. May 2010. Available at http://www.onderzoeksraad.nl/docs/rapporten/Rapport_TA_ENG_web.pdf

[Ricque 2009.1] Bertrand Ricque, Note to York Safety-Critical Systems Mailing List, 25 May 2009. Available at http://www.cs.york.ac.uk/hise/safety-critical-archive/2009/0166.html

[Ricque 2009.2] Bertrand Ricque, Note to York Safety-Critical Systems Mailing List, 26 May 2009. Available at http://www.cs.york.ac.uk/hise/safety-critical-archive/2009/0168.html

[Ricque 2009.3] Bertrand Ricque, Note to York Safety-Critical Systems Mailing List, 29 May 2009. Available at http://www.cs.york.ac.uk/hise/safety-critical-archive/2009/0173.html

[Ricque 2009.4] Bertrand Ricque, Note to York Safety-Critical Systems Mailing List, 2 June 2009. Available at http://www.cs.york.ac.uk/hise/safety-critical-archive/2009/0179.html

[Spiker 2009] Rolf Spiker, Note to York Safety-Critical Systems Mailing List, 29 May 2009. Available at http://www.cs.york.ac.uk/hise/safety-critical-archive/2009/0174.html