UNITED STATES OF AMERICA

NUCLEAR REGULATORY COMMISSION

+ + + + +

34TH REGULATORY INFORMATION CONFERENCE (RIC)

+ + + + +

TECHNICAL SESSION - T7

HAZARD ANALYSIS FOR NUCLEAR AUTOMATION:

DEFEATING DIGITAL DEMONS

+ + + + +

TUESDAY,

MARCH 8, 2022

+ + + + +

The Technical Session met via Video-Teleconference, at 3:08 p.m. EST, Stephanie Coffin, Deputy Director, Office of Nuclear Regulatory Research, presiding.

PRESENT:

STEPHANIE COFFIN, Deputy Director, RES/NRC

SUSHIL BIRLA, Senior Technical Advisor for Digital I&C, Division of Engineering, RES/NRC

PAUL BUTCHART, Instrumentation and Control Engineer 4, NuScale

MATT GIBSON, Technical Executive, Electrical Power

Research Institute

SHEM MALMQUIST, Captain, Visiting Instructor, fellow Royal Aeronautical Society, Florida Institute of Technology

PAUL REBSTOCK, Senior Instrumentation and Control Engineer, Instrumentation Controls and Electrical Engineering Branch, Division of Engineering, RES/NRC

JOHN THOMAS, Director, Partnership for Systems Approaches to Safety, Massachusetts Institute of Technology

MARK VERNACCHIA, GM Technical Fellow, General Motors Company

ALAN WASSYNG, Professor, McMaster University

P R O C E E D I N G S

(3:08 p.m.)

MR. REBSTOCK: Greetings all. And welcome to our RIC session on the management of digital demons. I'm Paul Rebstock, and I proposed this session to address an important element in the assurance that digital technology employed in nuclear applications will indeed do what it's supposed to do.

Digital systems are radically different from the legacy systems that they replaced. Everybody knows that. They don't just do the same job differently. They have the ability to do radically different jobs, and to allow for, and benefit from, interactions among tasks that have traditionally been kept separate.

Hardware failures still exist, but they are no longer the major source of problems. Hazard analyses offer a way to assess the behavior of systems, to uncover subtle behaviors and interactions that can be significant in digital systems but are less important, or even impossible, in legacy systems. Those are the demons, a wily bunch of imps bent on causing chaos wherever they can.

Our panel of renowned experts will

explore the capabilities and limitations of hazard analysis in the assessment of digital technology for use in nuclear applications.

Now I will get out of the way and let our experts dive in.

Stephanie Coffin, our Deputy Director for the NRC Office of Nuclear Regulatory Research will take it from here.

Stephanie.

MS. COFFIN: All right. Thank you, Paul.

Can everybody, I'll just do a sound check, everybody hear me okay? All right.

And I have to laugh because we're a little late to the session because we were dealing with demons of our own on this side. So, very, very apropos the name of this session, Defeating Digital Demons.

So, I want to first start off by acknowledging and thanking our panelists who are bringing a wealth of experience to this session. And our moderator Sushil Birla will provide a brief bio of each of them as we enter into the moderated sessions.

So, you'll get to hear a little bit more

about their backgrounds.

Now, we're taking kind of an atypical approach for the format of this session. And so, can you go to the next slide.

There we go. And one more.

And so we're going to discuss this session through a series of three questions posed to our panelists and moderated by Sushil Birla. Dr. Birla is a senior level advisor here in the NRC in our Office of Research. So, there are not formal presentations, but a panelist is welcome to use visuals to support his response. And so, you might see some of that as part of the session.

So, I want to give you a little preview of the format we'll be using.

And so, can you click one more time.

So, what Sushil's going to do is present and explain Question 1 and direct it to one of the panelists. And then he's going to invite responses from all the other panelists.

And then after hearing from our panelists, we want to invite you, the audience, to participate. And so, as you're listening to this Question 1 and the panelists, you're welcome to

submit your questions on the panel on the right of your screen. And then towards the end, Sushil will select and sequence the questions that best fit the flow and the purpose of Round 1, Question 1.

Can you show that?

And so, and then following the audience questions -- you can go to the next one -- we want to have a polling question. And the idea is just a yes/no question. It's related to Round 1, and it's really to give feedback to the panelists on how well they're articulating their cases.

And so then we move on to Round 2. If you can do that.

The same sort of format: Sushil will pose a question, the panelists will address it. There will be an opportunity for audience Q&A. And then we'll have another poll.

And then we'll move into the final round, Round 3, same format: question to the panelists, invite responses, and engage the audience on some Q&A. And then a final poll.

And then after that -- if you can click one more time -- we'll have a formal closure of the session. We'll be mindful of the time and pay

attention. And if the time permits, we might be able to do some additional questions at the end of the session.

So, I think with that, I am going to stop sharing and I am going to turn it back over to Sushil.

Welcome our panelists. Glad to have you here.

DR. BIRLA: Thank you, Stephanie. Is my sound coming through okay?

MS. COFFIN: Yes.

DR. BIRLA: Great.

This session's based on the outcomes of a series of international workshops held in 2020 by researchers from Halden in Norway. That series of workshops was focused on understanding the state-of-the-art in the safety assurance of digital safety systems.

One of the limitations identified in those workshops was the ability to evaluate the hazard analysis of a digital safety system. This session drills down to understand those limitations through a sequence of three questions.

This is a research-oriented session. Mind you, there is nothing, no direct connection with

regulatory practices. We are here to learn about the state of knowledge.

The questions are focused on a small subset of the real world problems to make the research manageable and to fit within the 90-minute time limit of this session. And I've already lost 10 minutes out of that.

The context is set via this use case, digital upgrades of reactor protection systems for operating reactors.

The reactor protection function is a Boolean function. The logic is computationally much simpler than safety functions in the auto or aviation sectors.

The loss concern is the loss of the reactor protection function. And any condition that can lead to that loss is a hazard of concern.

The causes of concern are systemic causes, rather than random hardware failures, such as -- and I'll just go through this list quickly -- and any combination of relevant hazard analysis method is within the scope.

Within these confines, take a look at this first question. And I'll read it again later

more slowly.

The terms design diversity and hazard analysis, have different meanings for different people. We are going to limit the discussion of these terms for the limited scope of this session.

Design diversity for the purpose of this session is described through this example: Items A1 and A2, which may be two components or subsystems or systems, are performing the same function and receiving the same inputs. If both A1 and A2 are performing correctly, their outputs ought to be the same.

A1 and A2 are successful diverse systems if the same common cause does not degrade the performance of both A1 and A2. For example, the same latent design defect in both A1 and A2; some unwanted interaction between the item and its environment resulting from unexpected signal pathways, unexpected propagation through interconnections, some degradation through shared resources, such as shared computing resources or communication resources.

A1 does not degrade A2. A2 does not degrade A1.

And the panelists are welcome to add any

other constraints to limit the scope of their answers.

The hazard analysis of interest is shown using a reference model from IEEE Standard 1012.

The requirements from the plant level analysis drive the development and safety assurance of the system in three activity streams. The middle one in black is the development stream. The top one in green is the verification stream. And the bottom one in red is the safety engineering stream of activities.

Hazard analysis may be performed in each of these activity streams, but our focus is on the hazard analysis performed as a part of safety engineering -- the bottom red block.

Hazard analysis is performed at every phase of the development process, starting from the planning phase, continuing to the concept phase, analyzing the interactions between the system and its environment there, proceeding to the requirements phase, then to the architecture phase. And there may be many iterations in these phases before clearing the gate to advance to the next phase.

At each phase hazard analysis produces

the safety constraints or requirements for the next phase, and also the test cases for verification. Hazard analysis continues through the remaining phases: Detailed design, implementation, testing first the system itself and then the system integrated in its operating environment.

Now let us walk through Question 1 again to refresh your memory.

Can safety evaluation of a reactor protection system based on state-of-the-art methods for hazard analysis be as effective as the current practice based on design diversity?

Please be brief. Bottom line up front, ideally a yes/no answer to start with. I'll introduce each panelist when I invite his response.

And at this point I'm going to unshare this and invite a response from Matt Gibson. Matt, get ready while I find your slide here.

MR. GIBSON: What are you looking for, Sushil?

DR. BIRLA: The slide with which I want to introduce you.

MR. GIBSON: Okay.

DR. BIRLA: Showing me only a few windows.

It didn't show me the window I was looking for.

And now I am there. I can see all my windows. And here you are.

Matt Gibson is the technical executive at EPRI, 13 years there, leading the integrated digital systems engineering portfolio. A licensed control systems engineer, certified cyber security professional, he's got more than 40 years of broad digital I&C systems experience, including hazard analysis, digital architect role, design implementation and support functions, human factors engineering, and cyber security.

Matt, take it away.

MR. GIBSON: All right. Thank you, Sushil. Appreciate you inviting me. You should have my statements.

So, for Question 1, I don't want to disappoint you, but I don't have a yes/no answer for that. It's an interesting question. The answer, based on our research is yes, with some important clarification.

And the first one there are two clarifications I want to get to here. The first one is in a modern hazard and reliability analysis

combined with a systems engineering process we can indeed demonstrate a high safety assurance.

These structured methods can reach a safety and security and performance conclusion, create analytical evidence to support that conclusion. So, they're sort of capable of doing that.

So, our EPRI research on operational experience data and on these analytical methods and the combination of them help support these conclusions, which we've done over several years.

So, the goal of these methods is not to eliminate the function but, rather, to find the right combination of design characteristics within an external feed, the I&C function, that will deliver the needed safety, security, and functionality. It's a holistic thing.

This may include diversity at one or more levels of this facility design, as indicated by the analytical method. So, you know, visualize the composition/decomposition. Now, where do you put diversity, and why you put it there, and which part of the stack really makes a difference on the effectiveness and appropriateness of that?

The second clarification I want to get to is the expectation to use a comparable method. So, the expectation of comparability with existing methods shouldn't be the exact objective of looking at these new methods. Because these new methods are quite innovative in a lot of ways and should be judged on their own merits. Because they may be demonstrated to be superior to the current approach in all aspects of efficiency and safety. That's something we can't, we can't assume what we do now is the gold standard.

Deterministic, you know, diversity, you know, saying, hey, I'm going to be, I'm going to do these set things with a set redundancy is not necessarily a sound position. You know, we looked at a lot of research into that. And there's not a lot of real good experiments that support the validity, you know, that that really is a good idea in all circumstances. Certainly in a risk-informed area but not in all circumstances.

All right. So, what you can do is say it's got to be diverse a certain way and it's got to be redundant in a certain way, you now make that the design. And you may divert your systems engineering

analysis away from things that could result in emergent behavior later on when the system's running.

So, I think that, you know, your modern hazards and reliability analysis provide a better understanding to the safety function, and they can ensure safety and security in a more efficient and technically sound manner.

That's what I would say about it.

DR. BIRLA: Thank you, Matt. Thank you.

I'd like to ask Mark Vernacchia of General Motors. GM has a lot of experience with safety critical systems in vehicles. Any particular example that you can share with us after I introduce you?

Mark is a GM technical fellow. In GM a technical fellow is the highest position in a technical ladder in General Motors, comparable to the senior technical advisor position in the Federal Government.

He's the principal systems safety engineer for all GM propulsion systems worldwide, over 20 awarded patents, Master's in engineering sciences from RPI, Bachelor's in mechanical engineering from Purdue, both excellent engineering

schools. Professional engineer in the State of Michigan. Expert systems engineering professional recognized by INCOSE.

He chairs the Task Force for SAE-J-3187, recommended practice for applying system theoretic process analysis, or STPA, to automotive applications.

Mark.

MR. VERNACCHIA: Hello. Good afternoon, good evening, good morning depending on where you might be. And I'm Mark Vernacchia. And I have a similar answer to what Matt expressed.

I think it's a qualified yes for me. There are caveats along the way, as always. But one thing that I can talk about within the automotive industry is sometimes the idea of diversity or, basically redundancy, is not an option we have a lot. It's expensive to do things, and it's expensive to package and produce and implement redundant systems in all of the safety-critical systems that we have to manage.

So, one of the examples I have is looking at a -- and I'll share my screen here -- looking at -- I'll make it big like that and, hopefully, you can

see -- this is a component of an electronic throttle control device.

So, what happens is: in the old days that I can remember you used to have a pedal mechanical lever hooked up to the throttle itself. If you pressed the pedal, a mechanical linkage would open the butterfly valve and you'd go faster or slower, depending on how it is.

And so, we replaced that, oh, 25 years ago at General Motors, maybe a little longer, with a motor-controlled system. And you can read the text down below. I'm not going to read it for you. But it has a lot of advantages over a mechanical linkage. Gives us a lot more latitude with a lot more features that we can accommodate. So, in essence, it's throttle by wire.

Now, the issue is what do you do with a control system that misbehaves? So, should I put a second throttle body in there? Should I put a second motor in? Should I put a second throttle position sensor?

You know, in our world, every time we add something we also add the risk that that will fail, too. And so, what we do is we end up with something

else here -- we end up dealing with more of a watchdog approach that you'll see in the center.

The CPU's going to do the processing. The watchdog is a processor is an independent parallel device that will keep an eye on what the central processor is doing as far as its control of the motor. And, therefore, if we see something misbehaving or unexpected, the watchdog can actually pull the plug on the CPU and return us to an idle position.

And a very simple diagram looks like this to go through. So, we have pedal sensors. Something figures out what the pedal is doing. Something then says position the throttle this way. That's the request. The throttle execution goes out for a command.

But then we look at basically very simple things like the request was made and the position was commanded, how do those match up? If they agree within a certain tolerance, things are good. If not, the watchdog will shut it down.

And we've done millions of vehicles like this. And so, I think there's a fairly strong case that you can actually achieve the level of safety

integrity you need without necessarily having to have redundant systems. Just an example here.

So, I'll turn it over to the next presenter.

And you're on mute, Sushil.

DR. BIRLA: Thank you, Mark.

MR. VERNACCHIA: You're welcome.

DR. BIRLA: I'm going to request Captain Shem Malmquist. So, Captain Malmquist, let me introduce you.

Captain Shem Malmquist is a visiting instructor at the Florida Institute of Technology. Master's degree in human factors. Experienced, considered an expert in aircraft accident investigations. Active current pilot for Boeing-777s on international routes.

He's instructed in a variety of both internal aviation and transport aircraft. Numerous technical publications on flight safety and accident investigation. Automation and human factors lead for the Commercial Aviation Safety Team's Joint Safety Implementation Team's Loss of Control Working Group, and many such other bodies that investigate safety like the Aircraft Safety State Awareness Working

Group, Joint Implementation Measurement and Data Analysis Team.

He's a fellow of the Royal Aeronautical Society, member of the Flight Safety Foundation, and the SAE Flight Deck and Handling Quality Standards for Transportation -- Transport Aircraft Committee.

Shem.

CAPT. MALMQUIST: Yes. Thank you, Sushil, for saying that.

I don't have a slide for this particular item. But the issue of the potential for common cause failures, even though you have diversity, can be seen in examples as simple as the Boeing MAX experience that we had. There's been several other aircraft where the exact same thing was implemented where the systems were seen to be diverse.

The problem turned out to be in the requirements. And those were common throughout all the designs. So, even though you did have diversity in some of the designs in terms of redundancy, they all led to a common -- they all led to common, I guess you could say, fault modes that resulted in the accidents.

And a large part of that was that the

human reaction to the -- to what they were saying, to the complexity of the situation, was not fully accounted for in the standards. And so, we really ended up with some accidents that could have been avoided through a more robust hazard analysis method.

And diversity really had nothing to do with it. It didn't change the outcome at all.

I can't hear you. You're muted, Sushil.

DR. BIRLA: Thank you, Shem.

I'm going to request Paul Butchart.

Paul, I want to ask for your answer based on your experience after I introduce you.

Paul is a master's in computer science, software engineering. He's an I&C engineer at NuScale. Worked eight years there. And just built his I&C engineering skill on the foundation of three years as senior engineer performing safety evaluation for regulators, five years as control systems engineer at Idaho National Lab's Advanced Mixed Waste Facility, four years as test engineer at INL's Advanced Mixed Waste Facility, five years in the steel industry as an I&C engineer, and 15 years in the Navy as electronic technician reactor operator.

So, Paul, what's your answer based on

your experience?

MR. BUTCHART: Well, thank you, Sushil.

In my experience I believe it is technically feasible that we can use an advanced hazard analysis methodology in order to assist, if not completely eliminate, design diversity.

The design of the NuScale power reactor is such that we have, we had very limited space. It is a small modular reactor, a fraction of the size of the existing reactors. And as a result of that, we had to limit what implementation, what control devices we had.

And part of that, we used STPA to analyze our module protection system. And with the intent that the analysis we used we ignored all of the assistance of redundancy of diverse design, of single failure analysis to optimize each channel of, of our monitoring control functions.

And I believe we are very successful in that. We began our analysis during the planning stages, carried it through conceptual design, detail design, and development of our design solutions. And we continue to use it even now as we're going through the process of our standard design approval

submission.

And I've given presentations at the STPA conference at MIT where I was able to provide a little bit more data to support my claims.

And I do have to agree with Matt and Mark, it does have to be taken with, with some precautions. It's not a panacea. It's not going to, you know, immediately solve all the problems. But judicious application of proper hazard analysis can significantly reduce the resource requirements for diverse design diversity.

DR. BIRLA: Thank you, Paul.

I'm now going to call upon Professor Alan Wassyng. And let me search for the slide to introduce you.

Professor Wassyng has over 30 years of experience in safety critical software-intensive systems, including development, certification, industry-academia collaborative research in many application sectors: auto, medical, nuclear, particularly in Canada at the Darlington Nuclear Power Plant where digital safety systems were pioneered almost three decades ago.

He has run for over 15 years a software

consulting business, over 19 years as a professor at McMaster University in Canada where in the past he has been the Director of the McMaster Center for Software Certification and Director for the Software Quality Research Lab.

He is the founder and the first director of the Software Certification Consortium.

Alan.

DR. WASSYNG: So, thanks very much. Just getting my screen.

So, thanks, Sushil, and thank you for the invitation.

So, I listened with interest. But I think I'd already guessed what people were going to say. So, I'm not sure if you can see the slides.

So, I need slides for one reason mainly: to remind me what I was going to say. So, hopefully, it will also help people look at what I'm saying.

So, my straight answer is a little bit different. I think I'm going to be the only one here that says no outright. Because I didn't give myself the chance to say sometimes yes and sometimes no. Because as soon as I get into that situation I really want to say no.

And I don't have anything bad to say about modern hazard analysis methods. I think everything I've heard from the panelists has been true. There have been fantastic improvements in what we can do in hazard analysis. I just didn't see that that was the solution to the problem.

Even if I have seen the perfect hazard analysis, I think there is a role for design diversity to play in being able to show a safe system. And a lot of my research is related to safety assurance, producing assurance cases. And trying to do that without the design diversity turns out to be quite difficult. In fact, so far we haven't done it.

So, to tell you where I'm coming from in terms of assumptions and experience, so when I do safety assurance, or when anyone does it, I think it's most effective -- and I've heard from the panelists that they were thinking the same thing -- it's done before and during development for the system, not after the development. It's done for it to be safe.

So, you use the safety assurance to sort of evaluate the development of the system as well as the outcome. And what we are interested in is

people, process and product. And showing that all hazards were eliminated or adequately mitigated is just not enough to show that the system is safe.

One of the things that we found in practice is you need to show that the system delivers the behavior that was expected. If it doesn't deliver the behavior that was expected, then people find work-arounds. And those work-arounds quite often disturb the safety that we have carefully built into it.

So, why do we want design diversity? And I think people have already mentioned it. We avoid systemic failures, or we hope that we can avoid systemic failures in constructing and implementing safety.

So, what I'd like to do is have a look at something that was done on Darlington. Sushil mentioned Darlington. This is what we did:

We actually did have two diverse systems. But related to what I think that someone said earlier, I think it was Shem, this started at the very top level of doing the development. So, right from the goal level requirements we started introducing diversity.

So, in fact, why do we want to do this?

It turns out that we make mistakes when there is complexity in the systems. And what we can do is try and use diversity to add complexity in different parts of the systems. And I don't have time to go into detail of that, but when we planned this we actually had different complexity in different parts of these systems. And the reason behind that is there was less chance then of having systemic errors that linked both systems.

So, why do we want design diversity? So, we have quite a lot of anecdotal evidence and very little experimental evidence. There is no way I can sit here and say I can prove that it's necessary. I also haven't managed to prove it is not necessary.

So, in a nutshell, my question is would it be responsible not to use it when we don't have something definitive in terms of an answer?

And I'll stop there, Sushil.

DR. BIRLA: Thank you, Alan.

I'm going to request Dr. John Thomas next. And let me introduce John.

He's the executive director at the MIT Partnership for Systems Approaches to Safety and

Security, and the MIT Safety and Cyber Security Research. He is the co-director, along with Professor Nancy Leveson, of the MIT Engineering Systems Lab.

And he has affiliations with other labs, a very interdisciplinary person: safety and security labs, systems engineering, software engineering, complex systems.

He has, again, multi-disciplinary teaching interests: systems engineering, which includes safety critical systems, cyber physical systems, automation and control systems, requirements engineering, and human-centered software engineering which includes user interface design.

John has participated in research with industry, collaborative research in many sectors: automotive, aviation, outer space, and nuclear. And in some of these application sectors he's also contributed to the development of standards there.

John.

DR. THOMAS: All right, thank you.

Well, my short answer is yes. I think that we can use state-of-the-art hazard analysis methods to provide a unique benefit that is not

provided by design diversity alone.

Now, I think I actually agreed with a lot of the previous presenter who said no. But I was reading the question a little differently.

When I say yes, and when I listened to all the other presenters say yes, I did not hear an argument to eliminate design diversity or things like redundancy. I think that you're going to be hard pressed to find someone who argues that we can eliminate across-the-board design diversity and it does not play a role in, in safety and what we need to achieve.

But there is a problem. I also think there is sound evidence that design diversity is not enough. So it's not necessarily an either/or, pick an apple or an orange. I think there is a role for both, both of these.

Let me show you on the screen.

Bottom line up front, some state-of-the-art hazard analysis methods have been proven to identify systemic hazards. And some are better than others. We're talking about things like engineering deficiencies, human interactions that aren't otherwise perceived and identified and addressed,

flawed requirements that aren't otherwise detected and prevented, complex interactions that aren't obvious to you.

This is a unique domain of state-of-the-art hazard analysis methods. And design diversity has not been shown to solve this satisfactorily, as I think all of the presenters would agree.

Ten years of evaluations and trials in the nuclear industry have demonstrated benefits from state-of-the-art hazard analysis methods for these areas.

Now, I'm using the word "some" here because I don't want to be misunderstood. There are a lot of hazard analysis methods. Some are better than others. Some have more evidence than others. I'm not talking about all. I can't make that blanket statement.

Also, an example from the software development on the space shuttle. They actually used a similar approach, where was this, 40 years ago to what's being proposed today in nuclear. They were 40 years ahead of the game at NASA.

They used multiple redundant computers, a backup flight system that was independent from four

redundant primary flights systems with similar functionalities to get us home if we need to. So, we've got five computers, one at least diversity on four. We're talking a different software development environment, different programmers, different contractors in case culture has something to do with it, different development advisors, different configuration and energy system across the board.

And the very first launch had to be scrapped because the backup computer, which was developed by Team A, couldn't be synchronized with the primary computer because they had different assumptions underlying them, and different requirements. And you couldn't figure out which one to believe and the whole thing had to be scratched.

So, redundancy and diversity I don't think is a silver bullet. It does have a role to play. So has complexity. And complexity is almost the root of the systems problem.

This is known as software diversity. It's also known as N-version programming. And we have a very rich scientific literature. This started, actually, in the '70s, and it's been studied very hard in the '80s and '90s. So, we're talking a

30, 40, 50 year old idea.

And that culture from the scientific studies cited at the bottom, the independence of assumption of errors as fundamental to the analysis of n-version programming, meaning software diversity, does not hold.

Why doesn't it hold? Because the independence is broken. Even if you've got separate teams, separate requirements behaving differently, we're making assumptions. Independent teams still share these biases, they share similar assumptions that humans tend to make, they share similar gaps in experiences.

The errors that we're seeing are not uniformly and randomly distributed. They tend to be clustered around common educations, common gaps in our experience of abnormal behaviors that we're not that familiar with, leading to common faults.

Here's a nuclear example. This is a real example. I scrubbed the details so you can't read this and figure out where this came from. But this is an event that happened at one of hundreds of events that have happened in the last ten years in the nuclear industry.

We've got two different supplier modules here highlighted in blue. Different suppliers, different requirements that they each created. It was seen as diverse according to the reviewers at the time where both were reviewed, both passed every check that we have. Independent requirements and implementation. It was tested and it was put into operation.

Months later -- and this is the nature of the problem, isn't it -- months later we get an interaction nobody tested for, nobody thought to check for. And both suppliers made this mistake. Both systems interacted in a brand new way that was overlooked. It happened in operation and led to a significant event.

Notably, neither one of these failed. So, we're going to go beyond a failure problem. There was not a single failure here. They worked as intended by each supplier. And that was the problem: the intention was wrong.

So, the question is can new state-of-the-art hazard analysis method reliably and consistently identify these flaws? STPA is one of the methods that have been tested on these cases - over a dozen

of these cases - from the nuclear industry. These are nuclear teams of engineers who learned the process, applied it, compared to previous teams who had applied other methods without the state-of-the-art hazard analysis. And the answer, yes, the empirical evidence exists. And the answer is, yes, these new methods are effective.

The old view is sort of these are random problems that happened. The world is full of only random problems. And as long as we have enough diversity, we're good.

I mean, that is part of the problem. But I strongly disagree that that's the whole problem. I think we've got very strong evidence that shows it's not the whole problem.

The systems view is it's not all about components and combinations of things randomly misbehaving, it's about the interactions between them, the common assumptions and how these have common causes between diverse equipment.

NuScale, on the call right now, Paul Butchart, did this in his company. This is data from his organization that he's presented. New hazard analysis method had some overlap in green for the

kind of things that we catch through diverse approach.

But we have brand new contributions to hazards and loss scenarios that are not identifiable and not preventable, as these numbers are, for the preventions that we're missing without a state-of-the-art hazard analysis method by the STPA.

This was submitted to the NRC, by the way, as a part of the safety assessment, and it was approved. Embraer did this evaluation. Bottom line is 44 percent of the results from hazard analysis overlapped with what they could identify today with traditional techniques. But 19 percent were identified earlier with hazard analysis, which is very important. And 37 percent were identified only with state-of-the-art hazard analysis methods, and overlooked with this assumption that diversity is going to solve a problem.

Nobody caught it anywhere until these events started to happen. There are safety standards that are accepted worldwide based on this concept of using state-of-the-art hazard analysis to complement the traditional redundancy and diversity approach. And it's necessary.

This is a picture from the standard ISO 21448 which shows this common engineering concept of unknown unknowns. It shows that you need hazard analysis to shift and minimize the amount of unknown unknowns. It's the unique domain of this hazard analysis method. You don't get that by throwing diversity at the problem.

With that, I'll end. Thank you.

DR. BIRLA: John.

DR. THOMAS: Yes?

DR. BIRLA: There's a question, and I'm going to read this, too, for everyone's benefit.

Diversity -- reading the question verbatim -- diversity is the measure to address all the things that you cannot think of. To say that you do not need diversity means that you think there's nothing you did not think of. But since all systems and how they're used evolve or change over time, is this not an impossible assertion?

John?

DR. THOMAS: Can you -- I'm parsing the question. What is the impossible assertion?

DR. BIRLA: Well, the question, the gentleman who sent the question has a view of

diversity. And that view is that it's the measure -- uses the word "measure," maybe not intended -- but diversity is the measure to address all the things that you cannot think of. To say that you do not need diversity means that you think there is nothing that you did not think of. But since all systems and how they are used evolve or change over time, is this not an impossible assertion?

So, the claim --

DR. THOMAS: The assertion is that, that we do not need diversity. I think nobody is saying that we can't, that diversity is irrelevant, diversity is -- that we don't want, for example, you know, diverse components in providing a function.

I think every industry has diversity. But every industry outside nuclear recognizes that diversity is only part of the problem. We do not have perfect measures of diversity, which is part of the problem.

We have a history of believing that these components are diverse. And then we have an event and discover we were wrong, they're not diverse. There was something in common that nobody knew about.

And the solution in other industries has

been hazard analysis. It's the solution to complement diversity. And you can think of it as --

DR. BIRLA: Thank you, John.

DR. THOMAS: -- as a way to --

DR. BIRLA: You make the main point that you aren't advocating total elimination.

I see that Matt would like to say something. Matt, please go ahead.

MR. GIBSON: Yes. Good discussion.

I do want to come back around to the hazard analysis and the use of it and diversity need a context. All right. So, when we're discussing hazard analysis we run the risk of discussing it way to narrowly.

All right. That's kind of maybe what we're hearing here. Because you really have to use the hazard analysis method in context. Because along with the risk analysis and the reliability analysis you want an overall engineering process. That's the only way you can use it.

Hazards analysis is diagnostic: it finds problems with the design. You then have to take that insight and change the design. If your design requires diversity to be adequately reliable, then so

be it. But it can be at different levels of the system stack.

You can have a mechanical thing over here that's not related to your system that provides the necessary reliability, whole system reliability to achieve its mission objective under all circumstances. We've got to get to a holistic engineering view, not a very narrow, you know, like this to this. And, you know, I just want us to think about that a little more broadly.

DR. BIRLA: Thank you. Thank you, Matt.

So, there was a few other questions from the audience. But they fall into the topic of the subsequent question. So, I'm going to defer answering those or bringing up those questions from the audience.

I'd like to now conclude the discussion on Question 1.

Stephanie, over to you for the first poll.

MS. COFFIN: Okay. So, I think we're going to, Spencer, show the poll. And the audience should see the poll.

Okay. And can you give us a sense for -

- we're asking the audience the similar question we asked the panel members. And so looks like 60/40 split.

Spencer, can you tell if the rate of return is done, is slowing down?

Okay, so let's close the poll. And, Sushil, you'll use that information as you move into Round 2.

And Matt has his hand raised.

MR. GIBSON: Just something to everybody keep in mind. When you're showing the slides there's about a 3 to 8 second delay between that and when the people see it. I'm getting texted while we're doing this. And they're saying, I'm not seeing this slide.

So, if we just bear that in mind when we're talking from our slides, maybe put the slide up a little bit or figure something out about that delay because sometimes we're talking way past the slides before people see them, I think. But, you know, good to note.

MS. COFFIN: Thank you, Matt. Good point.

All right, Sushil, are you ready for Round 2?

DR. BIRLA: So, can you give me the score?

I don't have that screen.

MS. COFFIN: It looks like -- oh, it's changing. It's going 50/50 now. Getting close to 50/50.

DR. BIRLA: Where did it end up?

MS. COFFIN: it looks like 54 percent yes and 46 percent no.

MR. VERNACCHIA: Yeah. I don't know how realtime it is because the yeses don't seem to be changing at all. So, sometimes it's good to close the poll and then put the results up.

MS. COFFIN: Sushil, maybe ready for Round 2?

DR. BIRLA: Yes.

So, apparently a good part of our audience is not as convinced as some of our panelists are. And would like to now give some more evidence to the audience through the second question and see if we can sway some minds.

Does sufficient scientific evidence exist to support the assertion that the hazard analysis can be evaluated independently with consistency for correctness and degree of completeness needed to avoid design diversity?

For example, do independent verification and validation methods exist?

So, from the audience during the discussion on the first question this was one of the questions that came from the audience: how do you perform an independent V&V to ascertain the adequacy of the quality of hazard analysis?

So, I'm going to request Paul if he has anything to add here, Paul Butchart?

MR. BUTCHART: Yes. Yes, I do, Sushil.

And kind of expanding on something that Dr. Thomas presented, I'd like to share the context of the information from NuScale that he showed.

When we performed our analysis, and in STPA, if you're not familiar, you analyze the functions of the system rather than the components of the system. And you look at, you know, what are the various ways a specific function can not perform what it's supposed to.

This screen shows an example of one of our safety constraints which was developed from the analysis. And what we did was we built a cross-reference that compared the results of our analysis with our existing requirements.

These particular ones were early on. It was functional requirements, essentially conceptual phase. And what we looked at, we looked at the different requirements, whether or not they contributed to satisfying the safety constraint. And then any evaluation that was necessary to clarify that contribution.

And the numbers we ended up with -- and Dr. Thomas showed a portion of this -- when we did our functional specifications, we had a very relatively low number of relationships between the safety constraints and our existing requirements. Very few that had more than 10 relationships. More so with 5 or less.

And there were 82 safety constraints that were identified that showed no relationships to any of our requirements.

We took the results of that and rolled it into the next phase of design, the detail design phase for our protection system. And when we re-performed the cross-reference we had a significantly larger number of relationships in total. Many more with multiple greater than 10 relationships.

We were down to 20 safety constraints

that had no direct relationships.

And, again, the number 75 for 5 or less, that's primarily because we improved that, those numbers to 10.

When we got to the design solution phrase, again a dramatic increase in the number of total relationships. The one number that stood out for many people was the fact that we still had 15 safety constraints that had no direct relationships. When we analyzed those 15, when we were performing the analysis we looked not only at the interrelationships between the various components within the system, but we also looked at the interactions with the operators.

And, obviously, we're not at the point where we can engineer operators. And so those 15 safety constraints ended up being administrative in nature. And they were incorporated, they'd been incorporated into our technical specifications operation procedures, abnormal operation procedures. And so, at least in our experience we found a very clear, very clear evidence that it can be shown in real numbers what the impact of hazard analysis is.

DR. BIRLA: So, that's good experiential

data, Paul.

The question really is if there is a third party evaluator or assessor or certifier, are there mature V&V methods for that third party to evaluate an organization's hazard analysis of a digital safety system? I'd like to ask any other panelist who might have anything to contribute to answer that question.

MR. BUTCHART: I believe I can address that one as well, Sushil.

We have within NuScale an independent verification design group. And they have been -- they're in the process now of performing their own hazard analysis of the reactor protection, module protection system.

And I have yet to see the results of it. I have not been involved in it at all. Which, you know, as a very, very interested party is rather frustrating. But I understand the need for independence.

And I feel that, you know, this could either, you know, an organization like this could either be within an organization, within General Motors, within Boeing, or whoever, or it could be an

independent third party organization. And I believe they would, you know, they would see similar results.

DR. BIRLA: So, your key idea is that they would perform their own independent hazard analysis.

MR. BUTCHART: Yes. That's correct.

DR. BIRLA: Okay, thank you.

Alan, you had your hand up and then you lowered it.

MR. WASSYNG: Well, it's not exactly on that topic. It's just related to the discussion before in terms of --

DR. BIRLA: Hold your horses on that. We'll get to the other --

MR. WASSYNG: I don't mean question 1. I mean what was -- Paul was saying before that.

DR. BIRLA: Okay.

MR. WASSYNG: So related to the question that you asked now, I personally don't have any problem with the experiments that are going on in terms of showing how effective STPA is, for instance. I think it's great.

Going back to what John said earlier, I thought that the question really was whether we needed both diversity and good hazard analysis. And

if you look at this, one of the great things that I think that I liked about STPA is that it shows me feature interactions which includes the environment. So Paul was talking about the operator. So what we've had is we've had good analysis for feature operations where the feature interactions are in the system that we're building. These are feature interactions outside the system we're building. And that's what causes a whole lot of our problems.

So the fact that STPA and maybe other hazard analyses I don't know about -- but the fact that we can do things with that is a step in -- a huge step in the right direction. But if I interpret your question slightly differently in terms of is there experimental evidence that just using STPA means I don't need diversity, I don't think there is any.

So I know that everyone -- like John says, that wasn't the question. That was the question you proposed.

DR. BIRLA: Okay. Yes. Fair enough. So here's the situation: You folks were so interesting and had so much to offer that we went way beyond our time budget on all three questions. So

I'm going to drop the third question and the third poll, but there was an audience question that came in that pertained to the third question, and that was directed to Paul, but I'm going to ask all of you to think about that question that came from the audience. And that's about expertise. What kind of expertise do you need to get this kind of quality? So Paul Butchart showed some data and some results that were impressive, but is that expertise replicate-able?

So who wants to take on that question?

CAPT. MALMQUIST: Well, I can start.

DR. BIRLA: Matt?

CAPT. MALMQUIST: Oh, go ahead.

DR. BIRLA: Matt, your eyebrows were looking -- I see John (audio interference) --

(Simultaneous speaking.)

DR. BIRLA: -- well Shem volunteered, so let's go to Shem first. Let's go with Shem first.

CAPT. MALMQUIST: Okay. Good. All right. So in answer to the question, it absolutely does require training and also facilitation particularly the first few times until people get used to it. I find working with various

organizations and also students that I've been working with that they'll -- depending on their background they have to be pulled back on track multiple times during the discussion. But after a while people start to get it. It just takes a bit of time.

And of course we're able to get really, really good results. I know that John Thomas can talk to that more, but even with people that don't have subject matter expertise it's really amazing. But of course subject matter expertise gains quite a bit. I mean there's no way to escape that.

But it's really -- one of the things that we're working on at Florida Tech is creating a certificate program in these topics. I know that was something you wanted me to talk (audio interference), but I can hold off on that and let some other people speak to this first.

You're muted.

PARTICIPANT: You're muted, Sushil.

CAPT. MALMQUIST: Sushil? Sushil, you're muted.

PARTICIPANT: You're muted.

CAPT. MALMQUIST: Sushil?

MS. COFFIN: Sushil, you're on mute.

DR. THOMAS: We can't hear you, Sushil.

DR. BIRLA: So, Matt, I want to come back to you.

MR. GIBSON: Okay. I'll be (audio interference). I want to add to what Shem said.

DR. BIRLA: Okay.

MR. GIBSON: We have to be careful about the Maytag repairman problem. Okay? This is an organizational problem. You get people trained on STPA and the other parts that go with that because they need to be engineers, they need to be reliability analysts. They need all that, right? So if they don't do it regularly, they will atrophy. So it's not some -- like an organization wants to adopt this, you can't go out and say well I'm going to do STPA on this project and I'm not going to do it for (audio interference) a while. They have to create a center of excellence that does it regularly in order -- it's like playing baseball. We all know the rules, but we quit playing for two years -- golf, my favorite, I can go way up and down in that if I don't play regular. It's the same thing. So I just wanted to add that to the things we have to consider when we

adopt these advanced techniques.

DR. BIRLA: Thank you, Matt.

John, you wanted to say something?

DR. THOMAS: Yes, I do. Questions 2 and 3 are related in my mind. Question 2 is about evaluating the analysis. Can we evaluate analysis consistently? And question 3 is do we know what affects quality of the analysis? So they're both really connected in my mind and I want to address them together.

As we discuss, I am getting a new appreciation for the concepts here and especially appreciate Alan's clarification. That just was very, very helpful.

So I think we've got a sort of consensus that diversity and hazard analysis offer -- contribute unique insights. We need both. We can't get rid of them. We can't eliminate diversity across the board. But we also know we can't have diversity maybe on everything. I know in aircraft there are certain things physically you can't have diversity. You don't want two actuators moving the same surface of an airplane wing because that's just a recipe for trouble. So we're forced

to have one jackscrew in certain places and other things like that.

So here's the question that we're posed: It says needed to avoid design diversity. So looking at this slide with this example we know in hindsight their strategy to use design diversity alone to go to two different suppliers didn't help the problem. Both parties made the same assumption. Nobody caught it. Everyone believed it was diverse at the time, which is more important than hindsight saying, yes, but it wasn't really diverse. That doesn't matter if they have no way to figure that out at the time.

So on one hand before this problem causes an event what do we do? Do we throw diversity at it using the current approach or do we throw hazard analysis at this? Now I think the best answer is let's do both. But take that off the table and then we get to the heart of these questions. You can't do both. You got to choose one. Do we throw design diversity, the traditional technique for decades at this problem or do we use hazard analysis?

Well, in this case we have one -- we have both halves of the experiment conducted actually. We have a factual answer for this case, and for dozens

of others as well, but let's just talk about this one.

The factual answer is they threw design diversity at it, right? They threw that approach. They strive for design diversity. They evaluated design diversity. They gave it a green light and it didn't work. They also took a set of blind engineers and said let's evaluate the logic, the functionality of this thing, which by the way shared among both of these. They said let's just evaluate the functionality, the shared functionality, right?

You could implement that functionality in one module or in two, or three, or whatever. They said let's get the functionality right using STPA, hazard analysis. And what they showed is teams, using STPA to get one definition of the functionality right, were able to eliminate this problem using hazard analysis alone.

And so if the team -- let's suppose the team had had this choice on the table to put two diverse modules sort of blind or in the dark, just give it to two suppliers and hope there's no common assumptions. Or we have the choice to give them hazard analysis to do a directed targeted analysis of

flawed assumptions that you would make, whether we're talking one module or two, or whatever.

And what we found is the hazard analysis found all the problems that were encountered. All the problems were caught with hazard analysis. They could have in hindsight gotten it right the first time maybe without even paying for a supplier two because they had the definition of all the logic and functionality of supplier one that was necessary to get the job done.

So I think there is an argument that this was tested, the hazard analysis was tested blind. It caught the flaws. Diversity was tested in the natural experiment because they did it for real and it didn't catch the problem. So I think the answer is a resounding yes. And this is just one case.

Now part of this problem is how do you review those assessments? Someone does STPA, claims it's done, how do you know they did it right? Well, that's been tested as well as part of these blind trials that have been done across multiple nuclear utilities, multiple applications. You can see on the left hidden flaws that were identified.

What happened is all teams consistently

identified the exact I&C errors, the exact common cause errors and failures that were missed with a diversity approach, sort of a blind diversity approach. The unmitigated human errors (audio interference) across the board. And we're talking multiple teams independent of each other, different geographic locations, different backgrounds. PRA experts participated although they weren't using PRA; they just thought that way. Digital I&C engineers participated and other folks. And every one of these was able to show that they can reliably use these methods to catch the problems.

That doesn't mean perfection. When I read this question I don't read this question saying do we have scientific proof of perfection? We don't have that for any method including how we evaluate diversity. But relatively speaking, relative to other methods does sufficient scientific evidence exist to support the assertion that the quality of the hazard analysis and the conditions that affect it are known to get useful consistent results? Yes. That we can say yes. Does it pass the bar of providing these consistent results that are understood well enough to provide quality? Yes.

Yes.

DR. BIRLA: Thank you, John. There's another question from the audience. Is qualitative hazard analysis sufficient enough for the decision makings in designing diversity and redundancy? Will a quantitative approach like reliability and consequence analysis be useful to support the decision making? Who would like to take that on?

DR. THOMAS: That depends a lot on the decision. I would say it depends a lot on the decision. There are some decisions where a probabilistic assessment is uniquely positioned to give you a good answer, but not all decisions. (Audio interference) a lot of decisions don't get any answer from a probabilistic approach or they get exactly the wrong answer. And we have evidence. We have tons of OE in this industry that have shown that.

DR. BIRLA: Mark, you had your hand up?

MR. VERNACCHIA: Yes, thanks. I just wanted to comment on a couple things real quick. One of the things that we found and we use STPA within General Motors especially for human machine interactions, and interactions is the operative word. Not interfaces, but interactions. And something that

I believe Paul said, that you need to analyze the functions and not the component.

Within GM we have a gigantic history of eFMEAs as the way to demonstrate adequate design integrity, but those are all a very failure-based approach. And what happens is we tend to focus on the components, do those FMEAs, and then we come around and put all those parts on a table and wonder why they don't fit as a system. Or if they do fit, why all of a sudden new behaviors, emergent behaviors appear.

And so this idea of having a analysis technique like STPA, you want to call it your state of the art approach, is an excellent complement to that. We have some inertia within our culture that if I had the biggest lever in the world, I could never get rid of FMEAs within General Motors. And that's really not my objective. My objective is to supplement those existing well-known techniques with new techniques that can address a system-level approach to things early in a process when there are still a lot of unknowns. And so --

DR. BIRLA: Thank you, Mark. Thank you.

We are coming close to the end of our

allotted time, so I want to read one more question from the audience. It is really very similar to a previously-asked question, but I want to honor the person who asked the question by presenting it to you again just to make you aware how significant this question is to our audience.

What evidence would suffice to claim -- to satisfy the claim that the people performing the hazard analysis possess the right skills to produce correct and complete results?

Shem?

CAPT. MALMQUIST: I think that the -- again would come back to the need for training and center-ization. I did want to mention just very quickly that the use of the hazard analysis method will at times point you in the direction of redundancy. In fact that's part of what we're doing when we're using it for aircraft design. It's not like we're missing that. The hazard analysis is actually highlighting where we need to do that, but also showing us where we don't need to.

But in terms of the system or the training, it really is necessary. We are in the process -- we've been hung up due to COVID of working

on a certificate program at the Florida Institute of Technology, or Florida Tech, which is -- again, we've run into some administrative hurdles. But some of the planned certificate programs could include system safety engineering, (inaudible), the overview, the theoretical vision, CAST, or Causal Analysis using System Theory, STPA, and safety management system approach. And I'm not sure to what extent the nuclear industry is using a safety management system, but there is a lot that can be done with that as well as cybersecurity. So it's something that might be possibly a way forward for some of the people where they aren't able to enroll in a current master's or doctoral program.

DR. BIRLA: Thank you, Shem. That is the kind of information that question was looking for.

So that concludes the discussion on question 2. You have actually answered question 3 as a part of this discussion, Shem.

So, Stephanie, over to you for the second poll.

MS. COFFIN: Okay. Do we have time for a second poll? If you could roll that out.

DR. THOMAS: In the interest of time are we able to discuss while folks are filling out the poll?

MS. COFFIN: I think we just have a couple minutes left, so I just -- I'd like to just focus on -- I'm not seeing any results. I don't know if others can see, but it just -- there we go.

So to remind folks, can the requisite quality of hazard analysis be evaluated independently with consistency? So a positive response. So you have been compelling in your discussions today. I do see it moving back and forth, but I think we're pretty close.

MR. VERNACCHIA: Again, I don't see any yeses fluctuating, so it's an interesting phenomenon.

MS. COFFIN: Maybe, Mark, it's the people who believe it jump right in there and then it's the ones who are on the fence that are (audio interference) in.

So thanks very much for participating in the poll.

Sushil, you need to give me a sense for time, of we need to move on.

DR. BIRLA: Yes, we should proceed to

conclude now.

MS. COFFIN: I think so. I think so.

Mauricio, are you able to show that slide?

DR. THOMAS: I'll stop sharing my screen. What's up there is the exact evidence that's required by the government --

MS. COFFIN: Thank you, John.

DR. THOMAS: -- (audio interference) applications.

MS. COFFIN: All right. So thank you, everybody, for -- Mauricio Gutierrez and Sushil and Paul Rebstock for putting together the session.

And I want to thank our panelists, in particular Mark Vernacchia and Shem Malmquist, for bringing us your knowledge from experiences outside the nuclear application sector. We have much to learn from you.

Dr. John Thomas and Professor Alan Wassyng for bringing us knowledge from your research across diverse application sectors.

And Matt Gibson and Paul Butchart, thank you so much for being here and sharing your knowledge in the nuclear field.

And of course our RIC technical support staff and the conference organizers for making this session happen.

And thank you, audience, for your questions and for participating the polls.

And if you have any more insights that you want to share with us, please -- there's a email there for Paul Rebstock.

And with that, I shall declare this session closed. Thank you all again for a great session.

(Whereupon, the above-entitled matter went off the record at 4:28 p.m.)