Official Transcript of Proceedings

NUCLEAR REGULATORY COMMISSION

Title: Advisory Committee on Reactor Safeguards
Digital Instrumentation and Control
Systems Subcommittee

Docket Number: (not applicable)

Location: Rockville, Maryland

Date: Tuesday, June 14, 2005

Work Order No.: NRC-461 Pages 1-296

NEAL R. GROSS AND CO., INC.
Court Reporters and Transcribers
1323 Rhode Island Avenue, N.W.
Washington, D.C. 20005
(202) 234-4433

UNITED STATES OF AMERICA

NUCLEAR REGULATORY COMMISSION

+ + + + +

ADVISORY COMMITTEE ON REACTOR SAFEGUARDS

(ACRS)

DIGITAL INSTRUMENTATION AND CONTROL SYSTEMS

SUBCOMMITTEE

+ + + + +

TUESDAY,

JUNE 14, 2005

+ + + + + +

ROCKVILLE, MARYLAND

+ + + + +

The Subcommittee met at the Nuclear Regulatory Commission, Two White Flint North, Room T2B3, 11545 Rockville Pike, at 8:30 a.m., George E. Apostolakis, Chairman, presiding.

COMMITTEE MEMBERS:

GEORGE E. APOSTOLAKIS, Chairman

MARIO V. BONACA, Member

SERGIO B. GUARRO, Consultant

THOMAS S. KRESS, Member

JAMES D. WHITE, Consultant

ACRS STAFF PRESENT:

SAM DURAISWAMY

MICHAEL R. SNODDERLY

ERIC A. THORNSBURY

NRC STAFF PRESENT:

STEVEN A. ARNDT, RES

RICHARD BARRETT, RES

JOSE A. CALVO, NRR

NORBERT N. CARTE, RES

CHRIS GRIMES, NRR

WILLIAM E. KEMPER, RES

PAUL LOESER, NRR

EVANGELOS MARINOS, NRR

ROMAN SHAFFER, RES

GEORGE TARTAL, RES

MICHAEL E. WATERMAN, Sr., RES

ALSO PRESENT:

MING LI, University of Maryland

A-G-E-N-D-A

Opening Remarks and Objectives

G. Apostolakis, ACRS 3

W. Kemper RES 4

Reconciliation of Comments on Draft Research Plan

M. Waterman, RES 7

J. Calvo, NRR 65

Draft Review of Reg Guide 1.97

W. Kemper, RES 103

G. Tartal, RES 105

Software Quality Assurance (3.2)

W. Kemper, RES 125

Assessment of Software Quality (3.2.1)

N. Carte 142

M. Li, UMd 157

Digital System Dependability (3.2.2)

S. Arndt, RES 213

R. Shaffer, RES 215

Self-Testing Methods (3.2.3)

S. Arndt, RES 275

Risk Assessment of Digital Systems (3.3)

S. Arndt, RES 281

P-R-O-C-E-E-D-I-N-G-S

8:30 a.m.

CHAIRMAN APOSTOLAKIS: The meeting will now come to order. This is the first day of the meeting of the Advisory Committee on Reactor Safeguards Subcommittee on Digital Instrumentation and Control Systems.

I'm George Apostolakis, chairman of the subcommittee. Members in attendance are Mario Bonaca and Tom Kress. Also in attendance are two of our consultants, Dr. Sergio Guarro and Mr. James White.

The purpose of this meeting is to discuss the NRC staff's Draft Digital Systems Research Plan, the staff's approach to revising Regulatory Guide 1.97, and two specific research programs discussed in the plan, software quality assurance, and the risk assessment of digital systems. The subcommittee will gather information, analyze the relevant issues and facts, and formulate proposed positions and actions, as appropriate, for deliberation by the full committee.

Mike Snodderly is the designated federal official for this meeting. Eric Thornsbury is the cognizant staff engineer. The rules for participation in today's meeting have been announced as part of the

notice of this meeting previously published in the Federal Register on May 31, 2005. A transcript of the meeting is being kept, and will be made available as stated in the Federal Register notice. It is requested that speakers first identify themselves and speak with sufficient clarity and volume so that they can be readily heard. We have received no written comments or requests for time to make oral statements from members of the public regarding today's meeting. I should note that the staff briefed the full committee on May 6 of this year.

We will now proceed with the meeting, and I call upon Mr. William Kemper of the Office of Nuclear Regulatory Research to begin the presentations. Bill?

MR. KEMPER: Thank you George. My name is Bill Kemper. I'm the section chief of the Instrumentation and Control Engineering Section of the Office of Research. We have numerous topics to cover in the next day and a half, and we have several presenters of the material. There's an agenda floating around. I presume everybody has that.

So before we begin, since we have some new members on our staff, I thought it would be productive to introduce at least the members of our staff that

will be making presentations over the next day or so. So Mike Waterman is here who will start out the reconciliation of comments on the draft research plan. George Tartal is in the back there. George, will you stand up, please? George joined our section about a year ago from the industry. He'll be talking about Reg Guide 1.97.

We also have Steve Arndt. Everybody knows Steve, I'm sure, he's been around for awhile. Steve will be talking about two or three of the presentations. Norbert Carte back there. Norbert joined us about six months ago from the industry as well. Norbert will be talking about software quality. Is Dr. Ming Li here by any chance? I guess he hasn't joined us yet. Okay, he'll be here later, from the University of Maryland. Roman Shaffer should be -- there he is in the background. Roman will be talking about digital system dependability. And Todd Hilsmeier, is Todd here? Okay, great. Todd's going to be talking about, tomorrow, dependability and analysis of digital system failure data. And he has Mr. Chu with him from Brookhaven National Lab. And also we have Professor Tunc Aldemir from Ohio State who will be talking to us later also about his research and investigation of digital system failure

assessment methods.

So, as you say, we're here to brief the ACRS subcommittee on various topics contained within our new Draft Safety Systems Research Plan, which covers 2005 through 2009. We briefed the full ACRS committee of this plan in May, and subsequently we were asked to provide more information on the research plan to the I&C subcommittee. So that is what we're here to do. Research has been working proactively with our stakeholders in NRR, NSIR, and NMSS to improve the draft research plan. We also hope to work closely with ACRS to improve our research program itself.

We appreciate the fact that ACRS has formed a subcommittee to support this area, and we look forward to our interactions with you all. We hope that these briefings that we're going to provide to the ACRS and its subcommittee on the draft research plan will result in ACRS endorsement of the plan, for our updated program plan, just as you did for the previous program plan. So unless there's any questions, at this point I'd like to go ahead and get started with the first presentation with Mike Waterman.

MR. WATERMAN: Good morning. My name is

Mike Waterman. I'm with the Instrumentation and Control Section of the Engineering Research Applications Branch in the Division of Engineering Technology. Let's see hee, background. I was with NRR's I&C section for about 14 years, and then I joined Research about a year ago. And one of the tasks I was given was to try to put together a research plan.

We started the plan about last year. We solicited comments in December/January timeframe. We received the comments. We incorporated comments from three supported offices, the Office of Nuclear Security and Incident Response, the Office of Nuclear Materials Safety and Safeguards, and the Office of Nuclear Reactor Regulation. And so today I'm going to go over how we addressed those comments briefly. So with no further ado.

In this overview, just a brief summary of the NRC licensing bases combined with the NRC licensing process, specifically NRR, because that's where my experience comes from. Talk a little bit about our emphasis on improving communications, and we'll get into the comment disposition summary table, and disposition of comments, and a little bit of a summary.

Briefly, just summarizing, you'll see later on, we had 34 formal comments received from the offices of NRR, NMSS, and NSIR. Thirty-one of the 34 comments were incorporated into the research plan, and the remaining three comments addressed topics that are really outside the scope of the research plan, or just required nothing to be done to the research plan. The first of those comments dealt with a suggestion that we put metrics into the research plan to measure the effectiveness of the research projects relative to the NRC's strategic plan. The second comment involved incorporating human factors considerations in our PRAs. We thought that would probably be better suited for the Human Factors Branch to deal with that in their research plan. And the final comment was something about NRR SRP is considered sufficient guidance for the fuel cycle people in NMSS, and didn't know what to do with that, so we just, you know, let it ride.

CHAIRMAN APOSTOLAKIS: But this is an important comment, though, isn't it? I mean, I read some of the memos, well, all of them actually, from the various offices to you, and I guess they all feel that what they're doing now is sufficient.

MR. WATERMAN: Well, I'll get into that,

Dr. Apostolakis, as I go through the discussion.

CHAIRMAN APOSTOLAKIS: No, but I mean the way you dismissed this last sub-bullet, I don't know. You said `I don't know what to do with that.' I mean, that's a pretty serious comment. They're saying what we're doing is good enough. When you form a research plan, don't you have to take that into account?

MR. WATERMAN: Yes, sir, we do. And I'll talk about that as we go on, and you'll see how all that folds out.

MR. KEMPER: Yes, we're going to address that as a common theme through several of our presentations.

CHAIRMAN APOSTOLAKIS: All right.

MR. WATERMAN: Essentially what the NMSS comment was was that they're moving toward a more qualitative risk-informed review, similar to what the NRR SRP already has in it. And what we're trying to do is get more specific than just qualitative, `This is a swell system' or `This is a good enough system,' things like that. So I'll get into that in a minute, Doctor.

RES revised the research plan to reflect the need for additional information in several areas on the basis of communications with the supported

offices. The research plan will continue to be updated in response to communications with the supported offices as new needs are identified, and as research projects are completed. And what I mean by that is the research plan ought to be a living document, not something we do once every five years, and then five years later go back and revise it. It should be a document such that as research is completed, we pull that research project out of the Section 3 of the plan, if you will, and have an annex where we describe -- summarize the results of that research, so that if somebody picks up the research plan, not only do they see where we're at and where we're going, but they can also get a flavor for what we've done and where we've been. So that's our vision of what the research plan ought to be, is something that continues to change as situations change.

MEMBER WHITE: Excuse me, I'd like to ask a question. As I was reading your plan last night, I was myself wondering about metrics by which you would evaluate your research effectiveness. In your slide here you say that that's outside the scope of the research plan. Of course any plan should have metrics, or goals, or targets. So is there some other document then that I can look at to see how you are

measuring your research effectiveness? If it's not covered in this presentation?

MR. WATERMAN: Well, we have NRC internal reviews of programmatic effectiveness that we're already using in the various offices. And primarily the reason I didn't incorporate the metrics to evaluate research effectiveness of the research plan, if I got into a long, lengthy discussion about how each of these things would be measured, if we're using PART, which is the Office of Management and Budget procedure, or something like that, we sort of divert attention away from the research into more attention devoted to actually measuring research effectiveness relative to the strategic plan. So it might be a good topic for a supplementary document that we can use to evaluate our research effectiveness, but I don't know that it goes into the research.

MEMBER WHITE: I think I understand what you're saying, but from a technical point of view, you surely have technical goals by which you would do a self-assessment of how well you're doing relative to those technical goals. And is that part of the presentation, and if not is there another --

MR. WATERMAN: It's not part of this presentation at all.

MEMBER WHITE: Okay, thank you.

MR. WATERMAN: Briefly, the NRC licensing bases depend upon the Code of Federal Regulations, Commission policy statements, standard review plans, Branch technical positions, consensus standards, regulatory guides that endorse consensus standards and take other positions, topical reports, and research reports. Now, these sources of guidance and requirements identify the safety system attributes that must be reviewed, and provide guidance regarding minimum acceptable standards of performance and quality. In a way, these documents, if you will are similar to technical specifications, for those of you who are familiar with those, which identify limiting conditions for operation, action statements, set points, surveillance requirements, and technical bases. The acceptance criteria identified in NRC regulations, guidance, standards, and technical reports are similar to surveillance requirement acceptance criteria. For example, nuclear power plants have a tech spec surveillance requirement to perform a heat balance, if you will, and use the results of that heat balance to adjust their nuclear power range instrumentation. Now, nuclear power plant procedures, not the tech spec, specify how the heat

balance is to be obtained, how the result is to be compared to nuclear power range instrumentation, and how the adjustment of nuclear power range instruments is to be performed. And similar to technical specifications, NRC regulations, reg guides, standards, the SRP, technical reports, prescribe surveillance requirements, if you will, but generally do not provide specific procedures for performing those surveillances. A major focus of this research plan is to produce the supporting surveillance procedures which will augment and supplement our existing process. We're not trying to replace process. We're simply trying to augment and supplement those with actual procedures such that no matter who does the review, they follow the same process, step by step, as much as possible. And right now those step-by-step procedures just, you know, they aren't there. I can say that from 14 years' experience of doing this that generally I had what was called an NRC audit assistant tool which didn't provide procedures but at least it guided me in what questions to ask. What we're trying to do is to formalize that process a little bit more so that no matter who does the review we get the same result. And that we're reviewing all of the things that we

need to review.

CHAIRMAN APOSTOLAKIS: But again, the standard review plan doesn't do that?

MR. WATERMAN: No, sir, it doesn't. The standard review plan has guidance that says you should check the correctness of a system through the various lifecycle phases, but it doesn't really go into the details of what does that mean, "correctness", what actual process do you go through to come to the conclusion that yes, the system is correct enough. All it does is it gives guidance. It's great guidance. I worked on doing -- I worked on writing the standard review plan with Gary Johnson out of Lawrence Livermore National Lab, and if you talk to Gary, he'll say the same thing I'm doing. The standard review plan was never meant to be a review procedure. It was meant to put bullets up of things that ought to be checked. The intent back when we wrote that branch technical position was to follow it up with actually writing procedures that describe when we say "correctness" what does that mean, how do you go through the process of assessing correctness, robustness, completeness, understandability. All of those attributes that you find in HICB-14, the branch technical position. All that we're really trying to

do is to augment and supplement the guidance that's in the SRP such that a reviewer can pick up that guidance and procedures and go through it.

If you go out to the regions, you know, they have inspection procedures for everything they do, and they do it -- they have an inspection procedure for a reason. It's so that every inspector does exactly the same thing so that the results are consistent. So that's what we're trying to do is to supplement and augment our existing procedures, especially now that we have large systems coming in that are going to require a lot of effort to review.

The NRC licensing process, the regulations, guidance, standards, and technical reports identify several hundred important attributes and associated criteria that must be addressed appropriately for digital systems to be licensed for safety-related applications. The emphasis there is several hundred attributes. The purpose of conducting research is to investigate current and emerging methods and knowledge, and where appropriate to augment and supplement NRC processes to enable NRC staff to evaluate digital systems consistently and effectively. We're already doing an effective job of licensing these systems, but the systems are getting

bigger, and if we don't start proceduralizing this review process, it's going to slow us down, and there is a potential there for being inconsistent in our reviews.

Now, with regard to additional emphasis on communications, the research plan was revised to provide additional emphasis on development of research products, review procedures, tools, etcetera, that augment and supplement existing NRC review plans and processes as part of a general process improvement initiative. Also we provided additional emphasis on enabling communications between research and the supported offices during the initial stages of research project planning to identify specific research products that must be developed, and during performance of research to keep the supported offices informed on the progress of Research.

Now, meetings have been held with supported offices to describe the research plan. We had presentations for the Office of NSIR, the Office of NMSS. We offered to present the research plan to the Office of NRR. They elected to not receive a presentation. That was back in the December/January timeframe where we wanted to just roll it out ahead of time, say this is what it's got, what do you think.

Two of the offices elected to see that, and they provided their input to us, and the other office elected not to.

CHAIRMAN APOSTOLAKIS: So what's TAG? T-A-G?

MR. WATERMAN: Oh, I'm sorry. In the future, what we want to do is set up technical advisory group meetings with participants from each office so that we can identify issues that are coming up, get the ball rolling on starting to do research to address those issues, or perhaps one office has an issue that another office has already addressed.

MR. KEMPER: The intent here is the research plan does not have the specificity needed to really sit down and write a statement of work. So the idea is it would provide a framework, general areas of research and specific topics that we could agree -- come to a conceptual agreement on. And then we would form the TAG and really flesh out the details of the specific scope and the applicable agency areas that are applicable to that in a TAG environment before we kick off a new project.

CHAIRMAN APOSTOLAKIS: There would be a number of these advisory committees, or just one advisory committee?

MR. KEMPER: No, they would be periodic, but certainly ad hoc as needed, basically to initiate any new work.

CHAIRMAN APOSTOLAKIS: I mean, it would be one group that will have representatives from NRR, NMSS, and so on? Or you will have one group from NRR, one group with NMSS?

MR. KEMPER: We haven't fleshed that out completely yet, but my desire would be to have all three offices in one TAG.

CHAIRMAN APOSTOLAKIS: I think that's a good idea.

MR. KEMPER: But you know, it may be that some projects supply more to one office than the other two, so you know, they could spend some unnecessary time in meetings.

CHAIRMAN APOSTOLAKIS: Yes.

MR. KEMPER: So we'll have to work through that and see what's the best environment for that.

CHAIRMAN APOSTOLAKIS: All right.

MR. WATERMAN: Now, as an example of communicating, NRR identified an issue recently on the need for regulatory bases that specify appropriate system architectures for digital safety systems, and the impact of those architectures on defense-in-depth.

A preliminary discussion between NRR and Research has identified the basic needs. Discussions will refine the objectives of the research and identify the products to be produced. We have yet to do that. I'm working on presentations on how I'm incorporating comments right now. Once I get through that I can get back to that work. This issue will be incorporated into the research project that addresses diversity and defense-in-depth. I haven't quite rolled that into the research plan yet, but that will be. It's a very interesting project brought up by Paul Loeser, and NRR identified it. It's if somebody is proposing to incorporate an RPS and SFAS all in one same microprocessor, so your trip and your mitigation systems all in one processor. It's just like, that's like all of your eggs in one basket. The microprocessor hangs up, you've lost trip and mitigation for that channel. I don't know, there's just something that doesn't ring true about that. So Paul's identified that. He's concerned about it, and he and I will be working together to try to hammer that out and see what we can do with it.

Well, the following slides summarize the disposition of the 34 formal comments RES received from NRR, NMSS, and NSIR. These are the formal

comments. We're also working with NRR to get some of their informal comments incorporated into the research plan as much as possible. The comments range from general comments on the contents of the plan to recommendations for revisions, additions, and modifications of scope. We anticipate that additional research plan changes will be made as specific research project needs are identified in the future. Again, this living document concept of the plan ought to be flexible enough to incorporate new research into it to be revised on a periodic basis.

The next three slides will show you a table of how -- this just kind of gives you an overview of the extensiveness of the comments, and how we address those comments. I really don't want to get into any discussion on the format of the table, or anything like that. It's just to kind of give you a flavor for how extensive the comments were, and how we changed the research plan to address those comments. Again, 31 out of 34 of the comments were incorporated. The other three, just couldn't fit them into the plan, so. But none of the comments were rejected, incidentally.

CHAIRMAN APOSTOLAKIS: Did you get any input from the offices regarding prioritization?

MR. WATERMAN: No, sir.

CHAIRMAN APOSTOLAKIS: Do they feel that some of these are much more urgently needed than others?

MR. WATERMAN: No, sir, we didn't.

MR. KEMPER: No, we hope that a TAG environment will address that. That's when we can really get the stakeholders together, and we can discuss that priority.

CHAIRMAN APOSTOLAKIS: When will you start implementing this plan? Have you already started?

MR. KEMPER: Well, some of the projects are already in progress. Obviously, they're carried forth from the last research plan. And as resources become available, and the timing is right, then we'll convene a TAG and we'll start the next.

CHAIRMAN APOSTOLAKIS: Now, I have the impression, and I'm asking whether you feel the same way, that this is a fairly ambitious plan, and you probably won't have sufficient resources to do everything that is in it. So somehow you have to prioritize.

MR. KEMPER: That's correct.

CHAIRMAN APOSTOLAKIS: Maybe getting input from the offices as to their urgent needs, although

they don't seem to need anything, judging from what I've read. That probably would be a good input to your process.

MR. KEMPER: Thank you. That's a good comment. We did make an attempt to resource-load the research plan, if you will. If you look back in Section 4 of the document itself, it provides detailed schedules, if you will, and the priority for each one of them. So we took a swag at the priority, if you will, based on our own intuition. But you're right, we have to confirm that with our stakeholders as we get into the details of these projects.

MR. WATERMAN: And that will definitely require a TAG, because I'm sure there's competing resources going on there. So one office may feel their priorities are a little bit higher than another one's.

CHAIRMAN APOSTOLAKIS: Well, even within the topics that are of their concern, I mean they should still give you some idea as to what the priorities should be.

MR. WATERMAN: Yes, sir.

MR. KEMPER: Absolutely.

MR. WATERMAN: So in the table, the revised information means the existing discussion in

the research plan regarding background issues, etcetera, was correct. In other words, if there was something that was factually incorrect in there, somebody caught it, and we corrected that. The added information means -- in the next column means additional discussion or amplification of the existing discussion was provided to clarify. That's, if you will, a perfective change to the research plan. And the revised scope column means the proposed scope of the research was revised in response to supported office comments. Some places where we thought we had the right scope, somebody pointed out it's not the correct scope, so we changed the scope in the plan on the next revision of the plan to incorporate that comment.

The following slides briefly summarize the comments received from the three offices, and the disposition of the comments. These slides only summarize the formal comments we received. I'm very anxious to also incorporate any informal comments we receive, verbal or whatever, into the research plan to address issues that were not conveyed perhaps clearly enough.

CHAIRMAN APOSTOLAKIS: Is there a reason why there are informal comments in addition to the

formal comments?

MR. KEMPER: Well, we've had several meetings with our stakeholders, as I said, to flesh out the comments.

CHAIRMAN APOSTOLAKIS: These are what you get in the meeting?

MR. KEMPER: Exactly. So the dynamics in the meeting, it fleshes out additional issues, and we certainly want to, you know, embody all those into the research plan that we possibly can. So that's what we mean by that.

MR. WATERMAN: And that's part of that communications thing that I think is really important. If we're not talking to our customer, if you will, then we're not really supporting our customer the way we should be supporting them. So that communications perspective, I've been given the privilege of actually writing up the office letter on memorandum of understanding of Research between us and NRR in this case here. And I have some ideas for how to improve that so we have a much more formalized process of communicating, and working together, and developing projects together up front so that when we actually get into the research it's going down the road that our supported offices actually need it to go down.

CHAIRMAN APOSTOLAKIS: Well surely though -- this is not a comment that directly refers to you, but surely Research has done work for NRR in the past and developed plans. So there must be some sort of communications process in place. You're speaking as if there is nothing there.

MR. WATERMAN: No, no, no. It's not that. I'm interested in process improvement as much as possible. I was over in NRR for awhile, and there were some things that I thought might be better implemented, and I want to incorporate ideas of process improvement into our research program, and one of those process improvements is iproving communications with our customers.

Now, this is -- in the following slides the comments are addressed in the order of the research plan sections -- in other words, Section 3.1, Section 3.2, Section 2, whatever -- beginning with a general comment on this first slide, the progressing through each research program. Within the body of the slides, each comment is summarized as a major bullet, which would be that bullet up there in white. And the research action to address the comment is then summarized in subordinate bullets, which, like green right here.

Now, this first comment regards how the research plan should be related to the NRC's strategic objectives and supporting strategies. And so, "In Section 4 of the Research Plan, each research project is linked to specific NRC strategic plan supporting strategies for achieving the NRC goals of safety, security, openness, and effectiveness." The other goal was management, but I really had a hard time working these projects into management. An in-depth discussion relating each research project to corresponding strategic plan supporting strategies would have been repetitive and ultimately distracting when you've got 24 projects and you're saying the same thing over and over for each project. The tabular format in Section 4 was considered the best alternative for succinctly relating the strategic plan goals to the research projects. So that's the way we went. At one time I was going to try to roll in those supporting strategies for discussion in our NRC strategic plan document. I just, after about five or six of those projects I thought, gee, I keep saying the same thing over and over. So we just put it down there as identifying it by number, which you can then pick up the NRC strategic plan.

CHAIRMAN APOSTOLAKIS: Is there -- It

seems to me the hard part would be to take any one of the projects you are proposing and prove that it does not relate to the strategic plan. I mean, safety, security, effectiveness, and openness. Just about anything you say is related to one of those, so I don't understand this comment. It doesn't make sense to me.

MR. KEMPER: Well, we might have gone overboard, but we really try to put an effort into each project back in Section 4 of not only identifying the goal, but also the supported strategies. So, yes, might have overdone it, but we thought it was an effort well spent.

CHAIRMAN APOSTOLAKIS: Actually, the research plan, it seems to me what you really want to see is what the differential would be, what the improvement would be as a result of each project in safety area, security, and so on, not if they are related. I mean, they are related. We know that. These four objectives of the strategic plan are so broad that just about anything you want is related to those. But when you talk about research plan, you really want to know is it going to revolutionize one area, are we doing nothing there and we're going to know what to do, or as Mike said, we know that we have

to follow some high-level process, but the details are not there. This is really what's important, I think, for the research plan to make sure that the reader understands. The reason why we're proposing this project is because in this area we have this need, and that's how we're meeting it. At least that's my impression.

MR. WATERMAN: Yes.

CHAIRMAN APOSTOLAKIS: So.

MR. WATERMAN: Part of linking this to the NRC strategic plan was it's historically that's the way we've always done it in the past.

CHAIRMAN APOSTOLAKIS: I understand. I see the word "stakeholders" is not there. Now, is it openness? Was it replaced? There used to be "stakeholders" someplace. Public confidence. Public confidence is now openness.

MR. WATERMAN: Those are the title of, you know, the objective --

CHAIRMAN APOSTOLAKIS: I'm not asking you to revise that.

MR. KEMPER: Thank you.

CHAIRMAN APOSTOLAKIS: These are your boundary conditions.

MR. KEMPER: Thank you.

MR. WATERMAN: Now, Section 2, we changed some things in Section 2, which is Objective and Scope. The first comment was to schedule periodic formal briefings for the supported offices on the interim results and status of the tasks. Research is developing more formal processes to improve communications with the supported offices, for example by the creation of a Technical Advisory Group or Groups, project development meetings, project status reviews. One suggestion I have that we may incorporate is to take our monthly status letter reports that we get from our contractors and extract relevant information from those and send it via email to our technical monitors, just so they're kept apprised on a month-to-month basis of what the process -- what project is going on, and how the progress is on that project, and things like that. So those things, that's a good comment there, and it's one that I fully support.

The next comment is, "Advanced instrumentation and controls research would also be beneficial for existing plants undergoing digital retrofits." And that recommendation was incorporation in Section 2.2, and out in Section 3.6, which is the Advanced Reactor Section.

CHAIRMAN APOSTOLAKIS: When do you think we're going to see this revised plan? I don't think we have it.

MR. WATERMAN: That's a good question. I think we intend to have all the comments incorporated by the end of this month.

CHAIRMAN APOSTOLAKIS: Okay. So sometime in July maybe.

MEMBER KRESS: Yes, sometime in July.

MR. WATERMAN: Most of them have already been incorporated, but it's just, you know --

CHAIRMAN APOSTOLAKIS: That's fine.

MR. WATERMAN: And I'd also like to vet it with my supported offices before we send it out to make sure I got their comment correctly, and that I've met all of their concerns, obviously.

So anyway, on the second bullet there, these sections were revised to reflect the potential applicability of advanced reactor research products. It was just, I think, adding in a sentence or two on, you know, it could be useful for existing plants.

Then we got into Section 3.1, which is the System Aspects of Digital Technology. And the first comment was, "The justification of Section 3.1.1 is to `reduce licensing uncertainty.' And the justification

should be focused on safety, improved efficiency, effectiveness, and realism, or openness." And when I went back to look at it, I said yes, heck of a catch there. I incorporated that into Section 3.1. Additional focus was placed on safety, although, because licensing uncertainty is a key issue in the nuclear industry with regard to digital retrofits, the focus on reducing licensing uncertainty was retained in there.

MEMBER KRESS: It seems to me like reducing licensing uncertainty, it is kind of a focus on safety, and efficiency, and effectiveness. That's what you have to deal with.

MR. WATERMAN: That's correct, but I think the issue with reducing licensing uncertainty revolves around that producing of review procedures. So that when a licensee submits a report, they know how it's going to be reviewed step-wise.

MEMBER KRESS: I see.

MR. WATERMAN: So that, you know, right now, you know, one of the things a licensee or a vendor asks when they do their kick-off meeting, they come in and they present their topical report, or whatever they're proposing that they're thinking about implementing. One of their questions near the end of

the meeting is always who's going to do the review. Now why would they ask a question like who's going to do the review? What difference does it make, right? Well, the reason they ask that is they know different reviewers have different slants on things, and they'd like to know what game they're going to be playing. So, you know, we're trying to reduce some of that uncertainty there. We'll all follow the regulations, but you know, some people are a little bit more tuned to one area than they are to another area. That's just human nature.

CHAIRMAN APOSTOLAKIS: Maybe some people are uncomfortable with the words "reduce licensing uncertainty". Maybe you can turn it to a more positive statement, and say "contribute to regulatory stability." Would that be better?

MR. WATERMAN: Yes.

MR. KEMPER: Sure.

MR. WATERMAN: I don't like to put negatives.

CHAIRMAN APOSTOLAKIS: Because you know, this implies there is now uncertainty, and why do you have uncertainty, this and that. Whereas if you say I want to improve stability, that's more positive.

MR. WATERMAN: Although there's an

implication there that we -- I know what you mean.

MR. KEMPER: Good comment, thank you.

MR. WATERMAN: Let's see. I'm going to get the transcript anyway, so I'll pick it up out of the transcript.

CHAIRMAN APOSTOLAKIS: Yes, we have a type of redundancy here. See both of you are taking notes, and there's going to be a transcript.

MR. WATERMAN: The next comment was, "The Research Plan and Statements of Work should include digital technology involving byproduct materials." When I went back through there, I realized, wow, I left a lot of our byproduct materials users out of the plan unintentionally. And so I incorporated, you know, `This research will support nuclear power plant licensing and byproduct materials users,' things like that. I did that in Sections 3.1.3, 3.1.6, 3.2, 3.3.2, and other sections as appropriate to bring that stakeholder more into the Research plan.

Now, "The state-of-the-art in software engineering may not be sufficiently matured for" and I put in brackets there "[quantitative] digital safety system reviews. This concern applies to the activities described in Sections 3.1.3, 3.2.1, 3.2.2, 3.3.4, and 3.6.3." And the recommendation was

incorporated. That statement was incorporated into those sections, and various methods will be validated as part of research and before recommendations are made to develop digital safety system review procedures. So the state-of-the-art may not be sufficiently matured, but that's what research is there to do, is to mature the process, and find out if that statement is in fact true.

CHAIRMAN APOSTOLAKIS: But I don't understand the meaning of this statement. It means the state-of-the-art is not sufficiently matured, therefore do nothing? Is that really the implication here?

MR. WATERMAN: Well, I didn't want to say that.

CHAIRMAN APOSTOLAKIS: This is probably the only comment that tells you that you need the plan.

MR. KEMPER: Well, I think the comment really was rooted in this. This technology may not be sufficient to implement these types of tools and processes that we're considering here. But as you say, it's -- that's exactly why we're doing --

CHAIRMAN APOSTOLAKIS: This is the only comment --

MR. KEMPER: -- and develop the technology, and you're going to hear many different versions of that in the next several presentations that we're going to make over the next day and a half.

CHAIRMAN APOSTOLAKIS: However, there is an implication perhaps that other people are developing the state-of-the-art, and all we do is take it and adapt it to our needs? I don't believe that. Because a lot of the models we're using were developed under the sponsorship of the Office of Research. Not out of the blue, of course. I mean, they are always building on existing methods, but this is really a strange comment. For the Research plan. It's a true statement, but for the Research plan it's a strange comment.

MR. WATERMAN: Well, it was a response to the Research plan from one of the supported offices. And we're working on that issue there, but you know, mind you, the comment was a lot bigger than this. And I think what Bill said was -- what the supported office was trying to say is that we're talking about going out and getting tools, for example. Well, how do we know the tools are even mature enough to do this. So, you know. And so that's part of our job is to find out.

CHAIRMAN APOSTOLAKIS: That's what Research is all about.

MR. WATERMAN: That's right.

CHAIRMAN APOSTOLAKIS: You make them mature.

MR. KEMPER: I think we're all --

CHAIRMAN APOSTOLAKIS: And besides, you know, we never rely on a single method in this agency. I mean, you know, quantitative methods may be one input to the integrated decision-making process. Words made famous by this agency.

MR. KEMPER: Exactly.

MR. WATERMAN: And the final comment in Section 3.1 dealt with Section 3.1.6. "Section 3.1.6 is not clear on how proprietary restrictions for `COTS operating systems' can be resolved in a way that can improve the assessment of digital systems." So Section 3.1.6 was revised to reflect that comment, that not all operating systems are proprietary, and to address issues regarding features of operating systems that may adversely affect safety. What we really want to know is for those operating systems you can look at, what things ought you to be looking for that could adversely affect safety such that you can bring it to the vendor's attention so that the vendor can correct

that potential safety issue.

And nuclear industry digital system developers have expressed a willingness to allow access to proprietary operating system design and development. The platform vendors have all done that. They've opened it up, and we review whatever we want to look at. It's when you get somebody like, say, an Allen-Bradley, a PLC goes in for a load sequencer, Allen-Bradley is a little bit more reluctant to allow us to peel back the lid, if you will. They have a small stake in the nuclear industry. They sell most of their stuff to much bigger customers. Dealing with those kinds of vendors is an issue, and I think that was probably the focus of this comment, was that when somebody is coming in with -- load sequencer is the one that comes to mind. People are going to digital load sequencers. They'll get an Allen-Bradley PLC, or Modicon, or something like that. And those vendors just, sometimes they don't want us looking at their operating system. That's proprietary information and they -- we have to do other things, like COTS-dedication process and things like that.

Now this first comment in Section 3.2 actually belongs in the next section on PRA. It just goes to show you how PRA can sneak into software

quality assurance issues. I'll address this comment in the next slide, so we'll start with the second bullet that says, "Link the objective of Section 3.2.3 to safety, improved efficiency, etc., and explain how NRC reviews can be improved to assess self-test features." Section 3.2.3 was lengthened to discuss the development of technical guidance regarding the use and review of self-testing features in digital safety systems. I suspect in future conversations we're going to have with our supported offices that section may be enhanced some more. What we're really trying to address here is, like operating systems, what features in self-testing do you need to look at, what features are appropriate for self-testing, and which features probably ought to not be used in self-testing.

My experience with the digital safety system failures that I've seen in the nuclear industry is it's always been self-testing features that have caused the cotton-picking failure. When we go out to review these systems, typically we don't have enough time to review every requirement in the system, so naturally we start by looking at the safety requirements, right? And we do our threat audits on safety requirements. Well, when you compare how much

software addresses safety systems, and how much addresses self-testing, you're like holy smokes. You've got this little bit of safety feature software, and this great big chunk of self-testing software that's supposed to make the product more reliable, and all the errors seem to be cropping up over in self-testing. So maybe we need some additional guidance on how to approach -- get our arms around that self-testing issue a little bit better.

The two failures I can think of that were caused by self-testing that I was directly involved in was the Turkey Point load sequencer. The self-testing feature locked out HPI in the system, with the intent that since it was continuous testing, it would only be locked out a little bit, and then if a signal came in, you know, nobody addressed what happens when a trip signal came in. That was one of those systems that it'd just stop the self-testing and start the process, as opposed to the approach that's now being taken by all of the vendors. And sure enough, the HPIs didn't get unlocked, and Turkey Point discovered that when Unit 4 was down, and one of their tests is to see if they can use Unit 3 HPI, and the crazy thing wouldn't start because the load sequencer wouldn't unlock.

CHAIRMAN APOSTOLAKIS: I read about it in

the report, and it was very, very interesting. And then the question came to my mind, to what extent is operating experience, nuclear and non-nuclear, driving the plan. Do we need to know that, or it's something -- it's just another project?

MR. KEMPER: Actually --

CHAIRMAN APOSTOLAKIS: This pointed out to me, you know, the real need of understanding the timing of things, and so on. So?

MR. WATERMAN: I think NRR's got a pretty good handle on the timing issues. I mean, when I reviewed the Siemens Teleperm XS, that was a big issue, was how are they timing all of this, what gets scheduled in for calculating trip, how do they schedule in the software testing stuff. Paul's done the same thing.

CHAIRMAN APOSTOLAKIS: But the question is broader though.

MR. KEMPER: Well, the use of operating experience for digital systems failures is certainly an essential element of trying to put together a priority system and specific tasking of the Research plan. Unfortunately, there's not a good user-friendly source, if you will, a readily available source of that information available to us. There's numerous

places, you know. We have our LER database, INPO has its EPIX, etcetera, etcetera. But when you go look at these information sources, it's very common that there's just not enough detail to fully understand and appreciate the mechanics of the failure itself. In fact, we've got a project which we've kicked off called the COMPSIS project. We're working with the Halden Reactor program to put together such a database, you know, with several international organizations participating for just this reason, so we can use it to better refine our research efforts in the deterministic world as well as the probabilistic world.

CHAIRMAN APOSTOLAKIS: Does the non-nuclear world have any general conclusions from their operating experience that we can take advantage of?

MR. KEMPER: Well, I believe that Todd will speak to that a little bit in his presentation tomorrow afternoon. That's one of the taskings in his project. But there are problems with that. I'll just kind of -- I don't want to steal too much of your thunder here, but different systems are qualified to different levels of quality, right? We in the nuclear industry of course set very high standards of quality, so when you try to compare failures of the same

platform being deployed across the process controls industry as a whole, it's difficult really to make a key judgment on the reliability of that equipment.

MEMBER WHITE: But what conclusion do you draw from that? Do you conclude therefore that you shouldn't look at that information, or just that it's hard to do?

MR. KEMPER: No. You should look at it, but you have to really evaluate it carefully to make sure you fully appreciate the ramifications of what you're seeing.

CHAIRMAN APOSTOLAKIS: We'll hear about it.

MR. KEMPER: Yes.

MR. WATERMAN: The other failure that I could think of is the ABB-Combustion Engineering developed an oscillation power range monitor for boiling water reactors. And that was a system that used master-slave microprocessors to check each other, make sure the channel was operable. And there was a problem on the 286 microprocessor chip that they were using with baton-passing. I don't want to get into a lot of detail on it, but what happened was because they had a slave processor, a self-testing feature if you will, the priority baton-passing down at the chip

level had a defect in it, and sometimes it wouldn't pass the priority baton back to the other features in that microprocessor, hung the microprocessor on a random basis depending on when you interrupted. And it took them about 10 months to work out that problem. That was all because they implemented a self-testing feature. So there's some issues with self-testing that we really need to get our arms around, and maybe do some more study on that.

In Section 3.3, which is Risk Assessment of Digital Systems --

CHAIRMAN APOSTOLAKIS: That's not a good title, is it? What do you mean by digital system? Building the hardware?

MR. WATERMAN: Yes, sir. It's hardware and software. It's not just software.

CHAIRMAN APOSTOLAKIS: But not the hardware -- not just the computers.

MR. WATERMAN: Well, it's not just the computers, that's right sir. For me a digital system is a system that consists of microprocessors supporting hardware, and the software integrated into that. It's not just software and hardware. It's the software integrated with the hardware.

CHAIRMAN APOSTOLAKIS: Some people might

argue that even for that digital system you shouldn't really talk about risk assessment, that you should talk -- I mean, if you want -- you should talk about a high-pressure injection system that utilizes digital technology and see then -- you do a risk assessment of the whole system, and eventually the whole plant. That prejudges what the -- I know that you don't have any ulterior motives behind this, but I'm just pointing out that there is some --

MR. WATERMAN: We know the device is digital safety systems, so I thought putting "safety" in there was kind of redundant. And I could have said "risk assessment of software and hardware, and software integrated with hardware" but for me "digital systems" pretty much wraps that up.

CHAIRMAN APOSTOLAKIS: Anyway, we'll see. We'll see --

MR. KEMPER: Steve is going to provide an overview later on today of what this is all about, this section of the plan.

MR. WATERMAN: So the first comment is "The plan should recognize that integrating digital systems into PRAs may not be practical and that a PRA may not be an efficient or accurate tool for digital system reviews." Of course, that's always one outcome

of your research. And we acknowledge that potential conclusion. We incorporate it into the plan that, you know, we may find out that PRAs are not the appropriate way to do it. But this issue ultimately will be addressed by the risk research projects.

The second bullet is "Include the integration of external events, environmental, and security issues unique to digital system risk into the discussion of PRAs." Section 3.3.2 was revised to state that these failure modes will be evaluated as part of the investigation of digital system failure assessment methods. However, the initial development efforts will exclude these external events, etc., until the methodology is sufficiently developed to address these additional issues. We're not just going to throw everything into the pot and then try to do one big research job with all of these different factors in there, you know. So small steps. Get to where you do something well, and incorporate the next issue.

The next two comments are, "The goal of the Section 3.3.3 research should be to provide methods for incorporating a digital component or system into a PRA. And in addition, acceptance guidelines should be considered as part of the

deliverable." And we agree with that, and we went ahead and incorporated those comments.

CHAIRMAN APOSTOLAKIS: What acceptance guidelines are these?

MR. WATERMAN: It's the acceptance guidelines for -- Steve can address that much better than I can.

MR. ARNDT: These would be issues such as what is the level of detail that you need for a system reliability model that includes digital components, what level of interactions between the process and between the various variables are necessary, if you're going to use the 1.7.4 criteria how do you interpret it for digital systems, or do you need to interpret it for digital systems.

CHAIRMAN APOSTOLAKIS: So you're really referring to the quality of the analysis?

MR. ARNDT: Yes.

CHAIRMAN APOSTOLAKIS: I think maybe you should use those words. Because acceptance guidelines usually means, you know, delta CDF.

MR. ARNDT: Yes, but there are other things included, like how do you interpret the defense-in-depth requirements in 1.7.4. But yes, we'll take that into consideration.

CHAIRMAN APOSTOLAKIS: It's important not to use a word for too many meanings -- with too many meanings.

MR. ARNDT: Okay.

MR. WATERMAN: So as I interpret really your comment, Professor Apostolakis, is we need to define what acceptance guidelines are.

CHAIRMAN APOSTOLAKIS: Well, use other words.

MR. WATERMAN: Flesh that out a little bit more.

CHAIRMAN APOSTOLAKIS: We don't need -- because usually, you know, in this context we mean guidelines regarding the acceptability of the change in terms of the risk metrics, or something else. Because the same thing applies to -- I mean, it's like Regulatory Guide 1.200, along those lines? What do we expect to see in the analysis?

MR. ARNDT: Yes. It's also along the lines, if you look at 1.75, 1.76, 1.77, those kinds of issues.

CHAIRMAN APOSTOLAKIS: Yes, yes, okay, good.

MR. WATERMAN: The next comment, "Section 3.3.3 should be clarified to reflect potential

capabilities and to ensure `risk' is not used in the plan as a synonym for `safety.'" And Section 3.3.3 was revised to reflect the comment, and then the Research plan was revised to ensure that the term "risk" is used where "risk" is required, and "safety" is used where the term "safety" is required. And there were places where that had to be changed.

The next comment, "Risk assessment should investigate advantages and disadvantages of analog and digital system architectures, and implementation characteristics in our PRAs." Section 3.3.4 was revised to include a discussion on evaluation of an analog Reactor Protection System, and an analog feedwater control system for comparison with equivalent digital systems to see what the delta was between looking at a PRA for your good old analog system, and how does a digital system change that PRA. So we've already got something in the shop for doing that, and we just needed to include that discussion in the plan. And so ongoing research is addressing the suggested approach.

And the last bullet in Section 3.3 is "Justify Section 3.3.4 statement that digital reliability assessment methods will reduce staff review effort by 20 to 30 percent." You know, I don't

where I got 20 to 30 percent to tell you the truth. We threw it out of there, took it out. At one time I thought that was a good number, but I was thinking about tools, and how much they might have been able to allow me to review so much more. And I came up with an estimate, but I took it out of there, because I really couldn't back it up by anything really hard and firm.

CHAIRMAN APOSTOLAKIS: Actually, they went on and said that in fact you may increase staff review effort. You remember that?

MR. WATERMAN: I would expect us to increase.

CHAIRMAN APOSTOLAKIS: At the beginning you should.

MR. WATERMAN: Yes.

CHAIRMAN APOSTOLAKIS: Because you're adding more. But an important element -- I mean, the staff review effort should not be the only metric here. We also want to do it right.

MR. WATERMAN: As a matter of fact, I don't think tools are ever going to replace the old eyeballs on the review. They'll augment. They'll do some things for us that maybe we couldn't do as fast, but when I went through reviews of a safety system,

and I was going through a threat audit, where I found most of the mistakes wasn't in like the design document, or the requirements document, it was at interface. All the mistakes start cropping up in those interfaces. How did you get from, you know, requirements to design. And I don't know of any tools that can actually pick that up. And sometimes, to tell you the truth, some of the problems I found, it was just a feeling I had when I reviewed it that something didn't seem right. I don't know a tool that's ever going to replace that, and when I dug deeper, I started uncovering, well, this is where they ran out of money on --

CHAIRMAN APOSTOLAKIS: Let me give you a little bit of advice here. When your contractors in the future come to you with Markov models, tell them what you just told us. And see how a Markov model can model that. I'll tell you, it can't. But I'm willing to listen.

MR. WATERMAN: of course, in the process of developing the model you learn something about the system.

Section 3.4, which is the Security Aspects of Digital Systems. We had some very good comments coming out of this. I'm still working with the Office

of NSIR to incorporate additional comments. This is a whole new issue for us, really. The first comment was, "Support development of 10 C.F.R. 73 requirements that implement NRC post-9/11 security-related orders and regulatory guidance." And that wasn't in the original security plan. That took Eric Lee working with me to help flesh that out, and we're working on that now.

The other bullet was "Support NSIR development of a comprehensive cyber security plan," and Eric and I are just now starting to work up the work breakdown structure on that. We had a couple of different ideas, and we need to hammer that down once I get off of the Research plan project.

"Section 3.4 should include research that supports industry implementation of NUREG/CR-6847, which is Cyber Security Self-Assessment Method. 6847, if you will, is similar -- when I read it, it impressed me as something very similar to a standard review plan, if you will. It identified things you needed to look at, and what was important, and those kind of things. But when it got right down to, well, how do I actually do that, it was like hmm. I don't know. Well, NSIR has stated that a tool is being developed outside through a multi-agency agreement, I

guess, that will implement the NUREG/CR-6847 guidance, and they'll use that tool on installations that already have networks. I don't know about -- I think we need to do some research on the networks that are being designed right now so we can catch problems early before they get installed into a plan. But we're focusing a lot more of our research on supporting this NUREG 6847 stuff, and I'll be rolling more of those comments into the plan as I get time before the end of the month, obviously.

Next comment was "Section 3.4.2 does not directly support NSIR plans, but it seems prudent to conduct research." This is on electromagnetic vulnerabilities, attack vulnerabilities. And "Though the Commission has not considered EM weapons as a credible threat to nuclear power facilities, some limited anticipatory research in this area is likely to be warranted." In other words, you know, as we find time, it's probably a low priority issue here. As we find time, we should be considering what do we do about low-energy radiofrequency attacks and high-energy radiofrequency attacks.

A related comment. "Section 3.4.2 describes an assessment of electromagnetic vulnerabilities. How does this activity relate to

TEMPEST programs?" TEMPEST is an acronym that came out in like the `60s and `70s. It's dated now. It's the Telecommunication Electronic Material Protected from Emanating Spurious Transmissions. And what TEMPEST really is designed to do is military and all the industries are now looking at, you know, people monitoring from a remote area, and picking up keyboard emanations, and things like that, and being able to take secure information out of a place by remote monitoring. That's what TEMPEST was designed to address, whereas what we're proposing in the research for electromagnetic attack vulnerabilities is completely different. I mean, instead of us worrying about what they're listening to, we're worried about what they're going to do to the instrumentation in the plant. That's the difference between those two. So apparently there was some misperceptions about what electromagnetic vulnerabilities involve, so I tried to clarify that in the Research plan with additional discussion.

And the next comment, "Wireless technology and firewalls should be subsets of a network security research project." That was a heck of a good comment, and so what I did was we used to have a Section 3.4.3 on wireless network security, and a Section 3.4.4 on

network security I believe it was, or something like that. Firewalls. And what I did was I combined those, per the recommendation, I combined those into a single section that is just titled network security. So Section 3.4.3 was renamed network security, and the discussion 3.4.4 was then just rolled up as a subset of that research. So that now the new focus, this is one of those revised scope things. The new focus of the new Section 3.4.3 is to address network security issues, including wire communications, wireless communications, and firewalls.

The next comment regarding security is "Section 3.4.3 should reference NUREG/CR-6847 which covers the assessment of wireless devices. The proposed research projects described should be informed with the assumption that licensees will implement the cyber security self-assessment tool described in the NUREG." And a related comment, "Firewall Security" -- remember, 3.4.4 is rolled up into 3.4.3 now -- "should state that the NUREG/CR-6847 can be applied to assess all digital devices, including firewalls, in nuclear power plants." I guess we'll wait and see how well the tool works out on that. "Revise the proposed research project to develop regulatory guidance on the use of firewalls

and expand review guidance to assist reviewers in evaluating the security risk of different firewalls." A terrific comment, and we're going to roll that into the Research plan also.

The Section 3.5, Emerging Digital Technology and Applications. "Discuss use of system diagnosis, prognosis, and online monitoring for virtual instrumentation and parameter estimation." And right now, the first version of the Research plan only talked about how it's being used for the diagnosis, prognosis, and stuff. And the comment was brought out that one of the other proposals for using this SDPM is to create virtual instrumentation where you use several different inputs to come up with a new output that could be calculated by it. And so Section 3.5.1 was revised to include a discussion on the advantages and disadvantages of using virtual instrumentation. The research objectives essentially remain the same because they were sort of generic objectives, keeping in mind that the purpose of the Research plan was to lay out broad areas, and then when we got into actual research projects we would nail down exactly what products had to be done. So throughout the plan we tried to keep the products generic enough that the plan remained usable for

whatever project we got into.

The next comment, "The regulatory applicability is not clear for the confirmatory studies of radiation-hardened integration circuits in Section 3.5.2." We've had discussions with the commenter on it -- point out that microprocessors, you know, the old let's radiation-harden it was let's hit it with everything we've got, good hard radiation, we'll see how well it works out. Now, some of the new microprocessors, they're kind of immune to the hard radiation, but if you put them under low dose, over time they kind of go to pieces. Kind of an interesting phenomena that they have more sensitivity to low dose rates than they have to high dose rates. I don't know the reasons for that, to tell you the truth, but you know. It is interesting. So when I brought that out, I think we're hammering that comment out. The tasks and products were revised to reflect the focus on guidance for the staff, and discussions with the supported offices, you know, as I say, we are clarifying that issue. You know, our old techniques of environmental qualification for radiation may need to be amplified somewhat to account for this low dose rate sensitivity.

And the next comment was -- this is all in

Emerging Digital Technology and Application -- "Application Specific Integrated Circuits and Field Programmable Gate Arrays described in Section 3.5.5 are not currently used in generically-qualified safety platforms." That comment was wrong, but the original comment before it was revised brought out that we've already reviewed some of this stuff. But all I had to go on was this comment until I actually talked to Paul Loeser and he showed me how it was misconstrued.

"Include, early on, an assessment of the existing or potential uses of this equipment in power reactors." The first paragraph was revised to reference current and future applications of ASICs and FPGAs. For example, I believe ASICs were used in the old Westinghouse 7300 Reactor Protection System. Westinghouse did a lot of work on Ovation. I think Eric Lee reviewed that when he was over in NRR. Ovation was an ASIC application. Toshiba I believe is coming in with field-programmable gate arrays platform applications. So the stuff is there, it's getting pretty close, and we probably should've started this research some time ago, but you know, nothing like now to get started.

Section 3.6, Advanced Nuclear Power Plant Systems. "Advanced instrumentation and controls

research would also be beneficial for existing plants undergoing digital retrofits." And that recommendation was incorporated in there. I added some additional words in there, but we addressed that earlier.

We had some general comments from NMSS fuel cycles people. "Review guidance in NRR SRP has been used recently by NMSS/FCSS for digital system reviews." Remember, I added that comment earlier, and Professor Apostolakis practically pointed out the unusualness of that comment. And so I revised Section 1.4 to state that NRC is conducting research to continually augment and supplement NRC capabilities. I can't emphasize that enough. We're augmenting and supplementing. We've got processes in-house. What we're trying to do is improve processes.

"NMSS/FCSS Regulations in 10 C.F.R. 70 are based on a risk-informed approach supported by qualitative acceptance criteria. Therefore, quantitative safety assessments and quantitative acceptance criteria may not be useful for the fuel cycle needs." And that's kind of strange. You know, it sort of sent me back. The Research plan projects in Section 3.3. address development of risk-based approaches for licensing digital safety systems. The

results of this research may support existing risk-informed licensing processes. But anytime you can move from qualitative to quantitative on your acceptance criteria you're taking a big step, in my opinion, toward improving your process. For me, qualitative acceptance criteria are, like I said, `this system is swell.' That's a qualitative assessment. `This is a great system,' that's another qualitative assessment. So I'd like to get us more toward a 95/95 type acceptance criteria, 95 percent confidence that it's 95 percent good.

MR. KEMPER: But I guess the key here is that our plan certainly has a risk component to it. And so we will look at fuel cycle facilities and see what we can do for them when that time comes.

CHAIRMAN APOSTOLAKIS: And we will pay attention to it.

MR. KEMPER: Absolutely.

MR. WATERMAN: NRR PRA boys had a general comment, or one person had a general comment. "The terms `software reliability' and `software quality' are used somewhat interchangeably." And the Research plan was revised to ensure there is a clear distinction between the use of the term "reliability" and the use of the term "quality." As I recall, we

have a project that used metrics, and what they're saying is they would use the metrics to predict reliability. And I think maybe you can use the metrics to predict quality, but I don't know about reliability. I don't know that quality and reliability are always directly related. I mean, you can have a quality system that doesn't do nearly what you want it to do, but it still works every time.

In summary, as I presented earlier, we had 34 comments from NRR, NMSS, and NSIR. Those were the formal comments. Thirty-one of the comments were incorporated into the Research plan. RES revised the Research plan to reflect the need for additional information in several areas on the basis of communications with the supported offices that I really would like to see continue. And the Research plan will continue to be updated in response to communications with the supported offices as new needs are identified and as research projects are completed. And that's the end of the presentation, Dr. Apostolakis.

CHAIRMAN APOSTOLAKIS: Thank you.

MR. WATERMAN: So we're working aggressively to incorporate the comments. Sometimes I've been known to lose my temper over being

frustrated, I can't get all the comments I want into the Research plan into the Research plan. But we're working on that issue. And I want the plan to be a good plan, no doubt about it. It also has to be flexible and adaptable. You know, who knows what the next issue coming up is, you know? If we were that smart, we wouldn't have any issues right now, would we? So it has to be flexible enough to accommodate that.

CHAIRMAN APOSTOLAKIS: Okay. Thank you very much. I see we have some extra time, so maybe we should invite other people to comment. Mr. Barrett first. Do you have anything to say on this, or do you want to add anything?

MR. BARRETT: No, I don't care to add anything at this point. Thank you, George.

CHAIRMAN APOSTOLAKIS: Thank you. Mr. Calvo?

MR. CALVO: Do you want me to do it from here or come to the table?

CHAIRMAN APOSTOLAKIS: It's up to you.

MR. CALVO: I'd like to come to the table because I think I need the overhead.

CHAIRMAN APOSTOLAKIS: That's fine.

MR. CALVO: Okay.

CHAIRMAN APOSTOLAKIS: Thank you, gentlemen. Appreciate it.

MR. CALVO: If you have no objection, I'd like Mr. Marinos and Mr. Loeser to join me at the table, if that's okay.

CHAIRMAN APOSTOLAKIS: Fine. How long is your presentation?

MR. CALVO: As long as you want it.

CHAIRMAN APOSTOLAKIS: No.

(Laughter)

CHAIRMAN APOSTOLAKIS: Well, actually, yes.

MR. CALVO: I'll tell you one thing. I'll send you the slides, of course the slides for the presentation, also for the backdrop slides. I went through the presentation. I cut out about five or six slides. So it's very short.

CHAIRMAN APOSTOLAKIS: Five or six is fine.

MR. CALVO: So actually, I will be addressing what we do. I'll be responding to some of the comments that Research has.

CHAIRMAN APOSTOLAKIS: So clearly identify yourself for the record.

MR. CALVO: Sure. My name is Jose Calvo.

I'm the branch chief of Electrical Instrumentation and Control Branch in the Office of Nuclear Reactor Regulation.

CHAIRMAN APOSTOLAKIS: And the other two gentlemen?

MR. MARINOS: My name is Evangelos Marinos. I was the section chief in the Electrical Instrumentation and Control Systems Branch. I was the section chief of the Instrumentation Section until May 16, when I was reassigned to a new position.

CHAIRMAN APOSTOLAKIS: Thank you.

MR. LOESER: My name is Paul Loeser. I'm a technical reviewer within the Instrumentation and Controls System, and at the moment, the remaining digital reviewer.

CHAIRMAN APOSTOLAKIS: Thank you very much. Okay, let's go on. Do we have copies of these slides?

MR. CALVO: Yes, you should have.

MR. SNODDERLY: George, what we'll do is these slides that are presented, we'll pass out to the members and to anyone.

CHAIRMAN APOSTOLAKIS: Well, we don't have --

MR. CALVO: You should have copies of

these slides because it was part of the package of the slides.

CHAIRMAN APOSTOLAKIS: Speak into the microphone, Mike.

MR. SNODDERLY: I said what I'd like is for you to present that material which you'd like to present, and then that would be publicly available. Right.

MR. CALVO: Okay. I guess what I'd like to do is what we do, what the NRR does. We've been doing that for several years. The staff reviews the process, not the product. And our process is contained in the standard review plan. They tell us how he's implementing the requirements -- not the requirement, the guidance of the criteria set forth in the standard review plan. So we leave it up to them. We don't tell them how to do it, we review what is there. And after we review the process, the lifecycle process, how we are putting a system together. We go back in for audits. We take a piece of the software, we go through it, and we determine how that thing is consistent with what they tell us. That's what we do. Now -- go ahead.

CHAIRMAN APOSTOLAKIS: You are telling us what you do. Is there an implication here that this

is good enough?

MR. CALVO: No.

CHAIRMAN APOSTOLAKIS: Okay. Let's go on then.

MR. CALVO: I'm saying this is what we do --

CHAIRMAN APOSTOLAKIS: No, that's fine, as long as we understand what you mean.

MR. CALVO: And now I'm going to tell you -- the next one will tell you what we have done. This is the systems that we have done. A Westinghouse. A more recent system was the Siemens, we have reviewed their platform. The Westinghouse also, ASIC. This is a functional modular implementation of a computer-based system. We issued a Common Q for Westinghouse on the combustion system, and that was Combustion, now Westinghouse has combusted together, and recently we have reviewed Triconex. We have reviewed the platform. We have reviewed the operating systems.

It's very interesting to note that the Siemens, the Westinghouse, and the -- wait, no, the Westinghouse and the Common Q, the operating system is not being developed in this country. It's developed by the Germans and the Belgians. Some kind of way the high level preparers are getting involved in the

operating systems, it was done in this country. So sometimes we had to go to Germans. We asked the Germans to come over here so we can ask some questions relevant to the operating system. So we have reviewed that the Triconex is the one that is actually located in Los Angeles, California, and the do that on their own. They have their own capability to do all these things. All the others, they don't have it. We invite them over, we ask them questions, but they're really platforms. Platforms tell you the operating systems, and we look for things like we don't like interruptions. We like for you to continue in a closed loop, which is normally about 50 milliseconds. It's a very simple system, the Reactor Protection System and the Engineered Safety Feature System. All you do, you go around for 50 milliseconds. And when you don't want to go, you hang around there. Don't go anywhere and come back, because you may not know where you left it, and then you get into problems. So it's a very simple system, very simple. The computers they use are the very lowest speed computer, because the lower the speed of the computer, the higher the reliability. So we're not talking about these 1 gigahertz. We're talking about 30 megahertz. 30 megahertz. They're very slow, and they're very

reliable. And you don't want to go up with it, because then it factors into the variability.

So this is the one that we have done. This is the one we plan in the future. HF Controls topical report. This is what is happening there. One thing I would like to bring out. I think the one that you're very much interested on getting involved is the Oconee. The Oconee challenged somewhat underlying principles and precepts of how you implement instrumentation and control systems, whether analog or digital. It's a very important one. The RPS, see we're thinking about the four echelons of defense-in-depth. We've got control systems, we've got protection systems, we've got engineered safety feature systems, and we've got display instrumentation. You've got the echelons that give you that kind of protection. What we want to be sure is that if one fails, you've still got the other three who are watching over that failure and can help you. In the Oconee, the combined are two echelons, but they combine protection and mitigation. And now we are concerned about that. Maybe we're going too far with that.

Now, I guess the question was asked today that -- by the way, Mike Waterman did a superb job.

A very positive presentation. I think I like the idea that maybe we're getting together now instead of moving apart. But I guess the question was asked, they almost asked you who is going to do the review. So some kind of way they figure out how they can get around it. They don't have to ask that question anymore because we've only got one left, you see. So, one question that we don't have to answer, all right?

Okay, that's fine. So the other one I'd like to show you is our perception of what we feel.

CHAIRMAN APOSTOLAKIS: Now, the Oconee license amendment request I bet is not risk-informed.

MR. MARINOS: No, it is not.

MR. CALVO: What?

MR. MARINOS: It is not risk-informed.

CHAIRMAN APOSTOLAKIS: It is not risk-informed, because we don't have any way of calculating --

MR. MARINOS: We're using the conventional approach that the Standard Review Plan guides us with to do the review as we have done for the other reviews that Mr. Calvo alluded to.

CHAIRMAN APOSTOLAKIS: Right.

MR. MARINOS: And this is a process that was developed with the assistance of the ACRS some

years back in a number of scientific institutions that helped us develop the approach that we have, which as Mr. Calvo indicated is a process-oriented approach for review.

CHAIRMAN APOSTOLAKIS: When was this developed?

MR. MARINOS: This was -- the final version of the standard review plan was issued in 1997. It started in 1993, if I'm correct, and in '97 it was published as a final approach for review. It was shared with a number of countries, in fact, the developed countries, England, France, Canada. And they gave us their advice, their guidance, and we developed that process.

CHAIRMAN APOSTOLAKIS: The reason why -- well, one major reason why it's not risk-informed is because we don't know how to do it.

MR. MARINOS: That's correct.

MR. CALVO: That's correct. Maybe one day in the future it will be defined. We're not there yet. We've got to --

CHAIRMAN APOSTOLAKIS: No, we will be there someday in the future if we don't keep saying we can't do it, let's not do anything about it.

MR. MARINOS: Additionally --

CHAIRMAN APOSTOLAKIS: I understand what your issues are. I mean, you have to make a decision within a reasonable amount of time, right?

MR. MARINOS: This standard review plan has not been fully tested, obviously, in this country as Mike alluded to, Mike Waterman. Duke Power Company's Oconee plant will be one of the best tests for us. However, the senior level scientists under their electrical instrumentation branch, which is a digital, he was assigned, in fact it was recommended by ACRS that he monitor the implementation of digital systems using the standard review plan at any other place where this is being done. And in fact, in Taiwan and in South Korea, they have implemented digital systems in the full scale, and our senior level scientist has monitored that, and the results are very positive in terms of guidance for doing the right thing. So this is what we base the --

CHAIRMAN APOSTOLAKIS: Yes. I'd like to know a little more about the Oconee proposal. And we can get the documents, I suppose, and have a look at them.

MR. MARINOS: The reviewer is Paul Loeser presently, so he can give you more details about the Oconee review.

CHAIRMAN APOSTOLAKIS: No, I'd like to get some documents first to read, and maybe in the future. But it's okay, there's nothing proprietary there or anything. I mean, you know, if there is we can look at it. So yes please, coordinate with Mr. Thornsbury.

MR. CALVO: The Oconee uses the Siemens.

CHAIRMAN APOSTOLAKIS: Yes. No, you said Framatome. Didn't you say Framatome?

MR. LOESER: Siemens sold that portion, the instrumentation section, to Framatome. When we started the review it was the Siemens TSX, now it's the Framatome TSX.

CHAIRMAN APOSTOLAKIS: Okay. But you had to go to Europe?

MR. MARINOS: Yes. Mike Waterman and myself and another employee went to Siemens to monitor there.

CHAIRMAN APOSTOLAKIS: The things one has to do. Okay. All right.

MR. CALVO: The board will view how we see the standard review plan. As you see, we have reviewed a lot. We have a challenge in the future. And what we're trying to do is trying to align ourselves with the Office of Research. We don't have enough researchers, and they don't have enough

researchers. And I don't know why we can't kiss and make up so we can all work together, with the goal of making the NRC look good at the end. I think we're almost there, okay? We have not kissed yet, but we're almost there.

CHAIRMAN APOSTOLAKIS: You have kissed, but you have not made up? Is that it?

(Laughter)

MR. CALVO: That's the toughest part.

CHAIRMAN APOSTOLAKIS: Well, your first bullet actually I think is great. I really would like to see that in every project. And that message will be sent loud and clear today and tomorrow. In each project, we want to know -- well, in different words, what are we doing now, what is the agency doing now, why there is a need for improvement, right? The problem to be solved, and how you're going to do it, how you're going to solve it. I think this is really the essence of the Research plan.

MR. MARINOS: We have gone through that, and Mike alluded to a TAG, the task action group, whatever.

MR. LOESER: Technical advisory group.

MR. MARINOS: Technical advisory group. And we did attempt this. In a previous attempt to

obtain concurrence from the NRR staff on the plan from 2000 to 2004. And we did periodically meet to discuss the various projects that they're proposing, and we didn't reach any conclusions of need on our part that they could convince us that it was there. So this is being proposed again, and I imagine maybe will be more successful.

CHAIRMAN APOSTOLAKIS: But what I'm saying is that I also subscribe to this kind of thinking, and we will -- and I'm sure the ACRS, judging from the way they reacted to the human performance research plan a few years ago, they think the same way.

MR. CALVO: If I may add, it's very important to know this, because we already review -- we only license a platform. We're going to be implementing about a hundred new plants in this country. If we're doing something wrong, we've got to know what it is before we can turn the wheels back. So that's important.

CHAIRMAN APOSTOLAKIS: I didn't get the impression from Mr. Waterman that the Research staff doesn't want to do this. I mean, this is a legitimate request. I mean, that's fine.

MEMBER BONACA: Yes. I see it more as a clarification. And really, for example, for the SRP

he provided an answer to that and said that they're developing procedures, and how to use criteria of the SRP in a consistent way.

MR. CALVO: Which is a good -- it's a good comment. And the reason for it, the standard review plan, whether you like it or don't like it, that's our bible, that's our criteria. They're mixing guidance and criteria in there. But we must move ahead with some instability in the process. If we're going to change it, why it needs to be changed, because we have a lot of trouble trying to convince the industry that you've got to change it for these reasons. It's going to cost you a lot of money and delays, and we'd like to know -- and that's the alignment that I'd like to have with Research in that area.

MEMBER BONACA: I didn't hear the word "change" in the issue of the SRP. I heard the issue developing a procedure to provide a consistent interpretation. So that could be useful to you, it seems to me.

MR. CALVO: That's fine. Which is a healthy review process, which is fine. I've got one more slide. The way we see what quality of research that we need from the standpoint of NRR. I'd like to give you a perspective of how we see the progress of

research. And this is the --

CHAIRMAN APOSTOLAKIS: Have you seen the movie "Dr. No"?

MR. CALVO: Yes, I did. James Bond. I don't know who won at the end, but -- Anyway, this is documented in all these non-concurrence memos that we have issued. It shows you the -- two or three of them, which I believe has something that we feel that has some value. But what is important here is not the memos. What is important here is the fact that yes, we've had meetings with Research, we have worked with them, and I estimate that when you do things at a working level and you start talking to each other, things get resolved. So we're saying here we've had a lot of meetings, and the project was discussed, but final version of the project has not been seen, and therefore may still not meet EEIB expectations. So we look like we're moving in the right kind of direction.

Now, there was a comment made that also, you say that informal comments were provided by the Research. So informal comments, it forces the staff to talk to each other, to align with each other. I'd like to propose that we had almost 18 projects that we have not discussed. Why don't we make them informal comments so we can talk about it, and the value of

those comments are incorporated into the program when they incorporate comments into the program. We're trying to be treated like the public. When the public provides you comments, we go through all the comments, and we resolve all the comments. We provide an answer to the public.

CHAIRMAN APOSTOLAKIS: But let me challenge you there a little bit, Mr. Calvo. I mean you are saying, for example, digital system -- 3.3.2, Digital System Failure Assessment Methods. And you say it's not desirable. Why isn't it desirable? How do you know it's not desirable?

MR. LOESER: The question we have here is what are we going to do with it. If we know --

CHAIRMAN APOSTOLAKIS: I didn't hear you.

MR. LOESER: What are we going to do with it from a regulatory basis? If we know that a particular digital system fails twice as often as different one, we can't tell the licensee not to use the one that fails more often. We can require them to take that into account. We can't --

CHAIRMAN APOSTOLAKIS: Is that what "failure assessment methods" means?

MR. LOESER: You said 3.3 --

CHAIRMAN APOSTOLAKIS: 2.

MR. LOESER: Oh, failure -- sorry, failure assessment. That has to do with setting up tools for assessing these methods. Once again, I wrote a couple of pages on the use of tools. I have some problems with the concept. If we make the use of a tool mandatory, then we are changing our regulatory method. If we make it advisory, what happens if the tool comes up with one result, and our conventional method of review comes up with another? Tools by their very nature become obsolete at the same rate as the types of things they are judging. If I have a tool to come up with the failure rate of a particular type of microprocessor, that tool is going to become obsolete as the microprocessor.

The biggest problem I had with all of these, however, is the way --

CHAIRMAN APOSTOLAKIS: Wait a minute, now. You're coming back again to reliability concepts, and this doesn't say that. This says methods of identifying system faults. So you're saying that methods for identifying system faults is not desirable by your branch.

MR. LOESER: No.

CHAIRMAN APOSTOLAKIS: I am not --

MR. LOESER: I didn't say that. What I'm

saying is the project, the way it was written with the intended results, or the intended products, and with the type of justification they have listed is not what we would have wanted. It was not discussed with us. We haven't had an opportunity to change it. In some of these instances where we said something was not required, not desirable, we have discussed this with Research. They have either been more specific on what they're really looking for. The one that comes to mind is the one on EMI testing. The project originally indicated they were going to throw open the entire issue of EMI testing, again which has been a number of times. It turns out what they wanted was there's one particular test that they think has a faulty premise. They have reason to believe this, and that's what they want to investigate. Once they stated it like that we agreed that this was a reasonable thing to do.

MR. CALVO: Keep in mind one thing. We never saw this research plan. We never saw it. We were not consulted to find out whether we align with each other. So when it's put on the table for us to review it, we had all those comments. This issue, they have discussed it with us, I think we can find a common ground. That's the big problem that we have.

CHAIRMAN APOSTOLAKIS: So what is -- I mean, the interactions that are happening now that Mr. Waterman talked about should have taken place before.

MR. CALVO: That's correct.

CHAIRMAN APOSTOLAKIS: That's a reasonable thing to say.

MR. CALVO: Agreed.

MR. LOESER: And I think if this particular project is modified, states what actually is going to happen, if we have some interaction I have no doubt we can come to some sort of agreement as to what should be done, why it should be done, and more importantly what the results are expected. When they state point blank that a reg guide, or a NUREG, on how this should be done, we question whether this is necessarily the right thing to do.

CHAIRMAN APOSTOLAKIS: There are two issues here, it seems to me. One is the view we have on the screen right now. And if I take the words literally, I don't understand why you fail to see how this would be useful. Okay? Methods for identifying system faults it seems to me would be useful to inform licensing systems. On the other hand, what you're saying is that the way the thing was written was not explicit as to what problem we're addressing, why that

is a problem, and how they expect to solve the problem in a way that would be useful to you.

MR. LOESER: Yes.

CHAIRMAN APOSTOLAKIS: And that has a lot of merit in it.

MR. LOESER: And I think we said that in each of our non-concurrences, where we stated that we think the solution to this is to get together with Research, discuss each one of these research plans, specify in a bit more detail exactly what they're after, what the products are. I think they should not make the assumption that it will necessarily, particularly when it comes to software metrics, or software PRA type issues. They should say that we will study this, present the reports, and then determine whether or not this should be turned into a NUREG.

MR. MARINOS: I'd like to make a last clarification with this language that is used there, system fault. We're not talking any actual physical system fault that they will identify. We're talking about ability to identify errors in the software that conventionally would not be identified by testing, or V&V, or this way. So certain tools are being proposed to be developed so that you can identify hidden errors

in the software, not system faults in the way that we interpret the actual physical system fault. We all need to know those if we can find them, but here is something that we're struggling. Research has tried to convince us that there is ways that we can find means by which we can identify those things, and then evaluate them. And as Paul alluded to, these tools that may be developed for a particular application, it will be actually for the same product if the software changes. Certainly it will be not available, and that will be usable for another product. So this is why we're relying on a process in developing those software, and of course, to complement this for security, we apply the defense-in-depth and diversity requirements, manual actions or automatic actions, to cover any uncertainty associated with software.

CHAIRMAN APOSTOLAKIS: No. I mean, you're absolutely correct, I mean there is -- we don't know what else to do, and we are doing the best we can. I mean, that's essentially that it is, diversity redundant. But let's not forget, though, that this is how the whole regulatory structure of the industry started 50 years ago, 40 years ago. And then with the advent of risk assessments, we found holes, we found improvements, and so on. And also, in all honesty to

this day, the system is still intact. I mean, we are risk-informing it, but not at a very high pace. Well, is it possible then that your traditional deterministic approach might have holes as well, and that if we try to do quantitatively, or develop methods for identifying faults, and go beyond that and do risk assessments, we may find holes. I mean, nobody's perfect, right? And the thing that I think -- don't you think you overreacted?

MR. CALVO: No, I'm not. I'll tell you what. I'm not.

CHAIRMAN APOSTOLAKIS: Look. It says not desirable.

MR. CALVO: Wait a minute. Wait a minute. Like I said before, we're moving ahead. We have reviewed and accepted many systems. And now, as we are responsible and accountable for the implementation of computer systems at nuclear power plants, I'm worried. I'm truly worried. Because there's nobody going behind me and helping me out to tell me you're moving in the right direction. I need that kind of support.

CHAIRMAN APOSTOLAKIS: Are you starting to get it now, do you think?

MR. CALVO: Well, I hope with your help

and this continued communication maybe we're going to get it. But I'm truly worried that we are moving ahead, and all we've got to do is get one system that fails because of the common mode failure. That's the end of the application of computer systems in nuclear power plants. We're going to put them on hold for a long time. And I need their help, but they've got to be focused on helping us out, to validate what we're doing, is it correct. You're right, we've got deterministic. I'm not quite sure if that's correct. I don't know the standard review plan gaps in there. We need them to focus and work with us, not to develop some new techniques and tools to do what? They all have been reviewed. There's nothing else to be reviewed at this time, only advanced reactors. That's something that you can put aside. They have limited resources like we do, and we need that help, we need alignment in here. It's very important.

CHAIRMAN APOSTOLAKIS: Well, it seems to me, Mr. Calvo, that your disagreement with the Research staff is more on the process that they're following to develop this research plan rather than the substance. You would like to see it more focused, which is legitimate, but you were really upset because you were not consulted before they put together the

first draft. But if one reads your comments, gets a very different impression, that nothing is of any value to you, and that is a little difficult to swallow.

MR. LOESER: First of all, that one column that said "desired by NRR" --

CHAIRMAN APOSTOLAKIS: EEIB.

MR. LOESER: It probably should have been re-termed as -- that we have a user need for it.

CHAIRMAN APOSTOLAKIS: Make it more technical.

MR. CALVO: Now, wait a minute, you're absolutely correct. This was a calling card. We need a calling card to put it on the table and tell Research, please, align with us and let's work together. That was the calling card. That was it. For an independent panel, you are looking at this, and decide, yes, it looks that way. But that was a calling card, let's start talking. And that was the whole purpose of it. Instead of start talking, it got worse, okay? And now we look like we are talking now.

CHAIRMAN APOSTOLAKIS: You are talking.

MR. CALVO: Yes.

MEMBER BONACA: Another I think really

actually a discussion between NRR and the Research on these issues by itself is going to improve the agency's capability, because there's going to be communication --

MR. CALVO: I agree.

MEMBER BONACA: -- and focus, and better understanding of what's needed and what's not needed. So I think --

MR. MARINOS: However, the process that we've had in communicating mutual needs is the user need, as Paul alluded to. So we had not expressed a user need because we were comfortable at least right now with the process we have in place through the standard review plan to do reviews. So when we were faced with this research plan, our concurrence, at least for the Electrical Instrumentation Control Branch would have been tantamount to a user need. And we said we have no user need, we don't need this research at this time. What we're doing is sufficient for us to convey to industry a coherent licensing approach. So that was the reason why we didn't concur as a branch on this program, because we had not identified a user need, and that is the only mechanism by which we would concur on a plan. So in an anticipatory research way, we wouldn't object, as you

are not objecting. You know, whatever they think they want to do. We don't want to second-guess them, but we certainly didn't want to be second-guessed either.

MEMBER BONACA: But in some cases, for example, you know, there has been today we are looking at PRA or risk evaluation as a fundamental support for fire analysis. And yet, there has been a lot of resistance in the past to developing risk-informed approaches to that.

MR. CALVO: That's fine.

MEMBER BONACA: Now, all I'm trying to say is that oftentimes, you know, you're looking at Research for more long-term, longer-term than you need instantly now. I think, you know, at that point communication is going to clear that issue. And you may agree that something can be done.

MR. CALVO: No, I don't disagree with you. I worry about that we move it ahead with a lot of reviews in here, with platform that we can review it, and I need help. I truly need help. This research program is looking from the researcher's standpoint, not from the agency's standpoint. And I just want to start getting together. The latest users needs that you had, which I think you had a copy of it, was in 2003. That established priorities, what you're going

to do first to help NRR to take care of its customers which are the licensees in this case. And I'm concerned. I'm truly concerned.

MEMBER BONACA: I must say, I'm pleased to see this move to yes, and not discussed --

MR. CALVO: I agree.

MEMBER BONACA: Because when I saw that the first time, reflecting on this, I thought that the "no" meant no need, desired no need, which is don't see any use for it. Now, this being converted into yes, with some changes, is beneficial.

MR. CALVO: Right. The "no" as presented indicated that we had trouble with it. When somebody hears you fresh, this is the program plan at Research, tell me what you think about it. So it was no communication. We just could not communicate even at that time, okay? We could not communicate. So we come out with the comments. And that was it.

CHAIRMAN APOSTOLAKIS: Okay. Now we have only a few minutes. Have you used all your view graphs or is there one more?

MR. CALVO: Almost done. I've got one more.

CHAIRMAN APOSTOLAKIS: Okay, one more.

MR. CALVO: I'd like to make some

suggestions for you to consider.

CHAIRMAN APOSTOLAKIS: Okay.

MR. CALVO: How we can go ahead with this. This is the latest users needs that we had, we prepared, we sent to Research. We need to update the old regulatory guides and go through because that has momentum. We'd like to bring them up to date, which I think Research is doing fine. And I think we can establish some priority which we want to see first. We don't want to review everything for the sake of reviewing it. We want to have certain things in there that we feel are important to our review process.

In state-of-the-art, monitor the cutting edge of what is done in other industries and academia. I think it's a good thing for Research. Keep abreast of what is going on out there, and maybe we can find out if something will have some implications on what we have done up to now.

The other one, new ways to regulate. At the moment these are primarily software-related.

CHAIRMAN APOSTOLAKIS: Let me understand this now, the second bullet, the state-of-the-art stuff. You would expect the Office of Research to produce some sort of a NUREG report, or some document that will summarize what is going on?

MR. MARINOS: Yes.

CHAIRMAN APOSTOLAKIS: And brief you on that?

MR. MARINOS: They have done so. And in fact one statement that was made in Mike's presentation that you commented on about not mature technology yet. It was actually right out of the NUREG that they produced and sent it to us for review about software reliability. And there was a statement there that the technology is not mature yet so we're going to back off a little bit and wait. So that's where the statement came from.

CHAIRMAN APOSTOLAKIS: So you appreciate this comment?

MR. MARINOS: Yes, we appreciate that.

MR. LOESER: This is I think Research Project 372.

CHAIRMAN APOSTOLAKIS: Who wrote that report, do you remember?

MR. MARINOS: Oak Ridge. I think it was Oak Ridge National Laboratory.

MR. LOESER: Actually, I thought it was University of Maryland.

MR. CALVO: I know we are running out of time. Let me go back, if you don't mind.

CHAIRMAN APOSTOLAKIS: Yes, go ahead.

MR. CALVO: New ways to regulate. We went to the software. It requires that when a method is discussed, we want to know the applicability of the method, what is the guidance. It's very important to distinguish what is guidance, what is criteria that should be used. And I think in our case, the method that we use is the standard review plan. Okay? Maybe somebody can help with this, pick up some gaps and holes in there, and maybe can identify those tools so we can do that.

The other point is how do we know that the method is properly applied, and that the licensee knows what he is doing? The acceptance criteria is needed. Okay, we're getting all this -- do you know how many it takes to review one of these systems? The platform? Something over one thousand hours. One thousand hours. And the criteria is about that high. And the guidance is about that high. That is a big help. We can focus on the important things. Help me. I need that help, okay? Right now we review everything, okay?

CHAIRMAN APOSTOLAKIS: Yes, but again, excuse me. There is a project somewhere here that says prioritize the thing using risk importance. Do

you say that's irrelevant?

MR. CALVO: No. Again, go back again, how that project was presented to us. All right?

CHAIRMAN APOSTOLAKIS: Okay, okay.

MR. CALVO: And then we go back again.

CHAIRMAN APOSTOLAKIS: We have settled that.

MR. CALVO: The "no" is not no, no, no. It's not ever no. It's tell me -- explain to me why, okay?

CHAIRMAN APOSTOLAKIS: Yes, okay. Fine, fine.

MR. CALVO: The other one is justification for the rejection of the license submittal if the quality is not present. What is missing, and what is important. We need that kind of help. Otherwise we're going to spend a tremendous amount of time trying to figure out that ourselves.

And I think the most important part, the most important part, for Research and NRR working level staff must work together to ensure that the application of the digital technology in nuclear power plants continues to be safe. And that is extremely important, okay?

Now, what I would like for ACRS to

consider, I recommend that the ACRS consider that all the internal staff comments on the research plan should be considered. All the comments. It's like the public comments. When you go for the communications to the public, you don't say `I got 50 comments from NEI,' all the others I don't care about. All the comments should be -- that would be the courteous thing to do. Review all the comments. You don't have to apply all the comments, but you learn something by the interchange. That's one thing I want the ACRS to think about that.

Then also, after review of the public comment, you recommend the disposition of the comments to be presented to the person who brought up the comments and to the ACRS. That's what you do when you've got the public comment. You come to the ACRS, and you discuss it, the public comment, and how do you resolve it. We want nothing else than that. We're not a second-class citizen. We're just like the public, American public, and we want to be treated like that. The only way we can be treated as public is to comment, and give you all those comments again? I think it's wrong, okay? And what I'm saying, we have not requested anything else that you have not readily provided to the public. And one thing I

recommended for you is to get involved with the Oconee application to replace the analog system. It's a very interesting application who challenges a lot of our principles and precepts. And brief the analog system, and why do you do things. In the analog existing and the digital system, they cannot be -- that's very important.

CHAIRMAN APOSTOLAKIS: So you want us to get involved in that?

MR. CALVO: Yes.

MEMBER BONACA: I would like very much -- I think it should be before the main committee.

CHAIRMAN APOSTOLAKIS: Me too. Me too. I was telling Eric here --

MR. CALVO: And another thing. We need your help on that one.

CHAIRMAN APOSTOLAKIS: Very good.

MR. CALVO: Because it's highly philosophical, broader, and we need that because it brings the whole aspect into that.

CHAIRMAN APOSTOLAKIS: Wonderful. So we can actually -- I mean, we can have the stuff.

MR. CALVO: Yes.

CHAIRMAN APOSTOLAKIS: Just tell us when will be an appropriate time to brief us.

MR. CALVO: So we can get Research to help us in this.

CHAIRMAN APOSTOLAKIS: Yes.

MR. CALVO: I think we both jointly can come in here and present.

CHAIRMAN APOSTOLAKIS: I really get the impression, I mean just to close this. You are really the decision-makers, right? You decide that something is acceptable to this agency or not. And you really want to know, if somebody says I'm going to help you, where he's going to help you, how he's going to help you, which point, you know. And this is a characteristic of decision-makers. I mean, you really don't want to see doing research for its own sake, and all that. So I see what the difference in approaches is.

MR. LOESER: I think research for its own sake is very good. But then it has to be presented as such, not as this is the solution to all your problems in five years.

CHAIRMAN APOSTOLAKIS: Anyway, no I think we understand, and the Research I'm sure understands.

MR. CALVO: I think you hit it right on the target, and that's what we need. It's very difficult for me to get a product from Research, and

then go back and look at the industry in the eyes and say `Hey fellows, I'm going to have to backfill you all this because of this.' I've got to give them the resource. If I don't have the resource, I'm going to be in trouble. Look, I want this very much.

MR. MARINOS: One last comment that you made about the regulatory uncertainty, and it was changed to regulatory instability. I think that the premise of the original statement was correct. I believe that this plan will create, and I've had already reaction from industry, it does create a regulatory uncertainty, because it places a cloud over the process we use and we have used to do major reviews. Those platforms that we've used are major things. And they're being implemented now to a plan which is equally challenging, but not as challenging as reviewing the platform. So how do you do this for the entire industry, for the entire world under this process, and yet we have this plan with 500 pages of tools by which they will second-guess the work that we do. That's where this regulatory uncertainty lies and it is, in my view.

CHAIRMAN APOSTOLAKIS: Okay.

MR. CALVO: Anyway, that completes my talk. Thank you for listening.

CHAIRMAN APOSTOLAKIS: Thank you very much gentlemen. Any comments from the staff?

MR. BARRETT: Yes. This is Richard Barrett, Office of Research. I'd just like to say a few things to clarify. First of all, we have a number of processes for gaining user office commitments to support our research program. And the TAG process is certainly one of them. The process we've used of developing this plan and submitting it for office concurrence is also a legitimate process. We don't always just sit and wait for a user need to come from the user office. This is an area where I think the Office of Research has justifiably taken the initiative to produce something that can be of use to the agency in the future. And I say that having recently come from NRR.

Also, I think it's not fair to characterize this as research for research's sake. I think what the Office of Research has done is put on the table a broad-ranging proposal. And we are open to technical comments. We're open to process comments. And we're anxious to work in a TAG environment with our user offices in the future. The Office of Research has a record of dealing openly with its users, and we will continue to act in that way.

CHAIRMAN APOSTOLAKIS: Okay.

MR. GRIMES: My name is Chris Grimes, and I'm the deputy director of the Division of Engineering in NRR. I want to clarify the point that Mr. Calvo described this as a non-concurrence, and that's true. The Office of NRR chose not to adopt all of the comments submitted by EEIB on the user need. While we do have an established protocol for the communication between the two offices, individual branches, even individual sections, tend to exercise the technical advisory groups to a greater or lesser extent. They have more or less effective communication between the two offices.

There has been an effort underway between the leadership teams and the two offices now for at least one year, maybe two, to try and have a more consistent treatment about user needs, and the reliance on technical advisory groups to coordinate the goods and services. And as you pointed out before, it's not sufficient to say that they are related to a strategic goal of safety, security, effectiveness, and efficiency or openness. The goods and services have to be related to how they contribute. In what way are they expanding knowledge so that we have a better understanding of safety, or

that we have a more efficient review process? I share Mr. Waterman's view that there ought to be a focus on process improvements and contributions. And to that extent, we felt that the majority of comments that were going to be proposed were not constructive, and that they would suggest the research plan should be -- the baby will be thrown out with the bathwater. So we only adopted those that we thought were constructive.

We do favor -- there is a consistent use of technical advisory groups on a regular basis. We will not wait from 2003 till 2005 to do the next comment or round of communications on the progress on the user needs, or any of the research plans. Our mutual offices will expect that a monitoring will be done at least on a quarterly basis, if not a monthly basis, to ensure effective communication.

MR. CALVO: If I may, a rebuttal, just a little bit. A rebuttal a little bit. Those comments that were selected to be given to Research that were NRR, they were never discussed with us. We don't know --

CHAIRMAN APOSTOLAKIS: Yes, this is internal to the office.

MR. CALVO: I know the communication problem is both vertical and horizontal. So we're

having that problem, not only between offices. It's within the office.

CHAIRMAN APOSTOLAKIS: Okay. Any other comments from the staff or members of the public on what we've heard? Well, thank you all. Thank you very much. And we'll recess until 10:40.

(Whereupon, the foregoing matter went off the record at 10:23 a.m. and went back on the record at 10:41 a.m.).

CHAIRMAN APOSTOLAKIS: Okay, we'll continue now with the revision of the regulatory guide, right?

MR. KEMPER: Yes, yes. If I could just --

CHAIRMAN APOSTOLAKIS: And Mr. Kemper, before George takes over. Go ahead.

MR. KEMPER: Thank you. I'd just like to make a few comments here. We're really here, George, at your invitation. This is a work in progress, and we're almost done with this reg guide, draft reg guide. But it hasn't quite gelled yet. So what we would like to do is to review this with the working group and get your comments. At this meeting, that's fine, or later on if you choose to write something and send it to us informally that'd be good too.

But basically the new reg guide endorses

a process which is a revision of the IEEE 497-2002 that's a new approach to identifying post-accident monitoring instrumentation. It uses a performance-based versus deterministic point of view. As you all know I'm sure, the current revision of Reg Guide 1.97 is very prescriptive. It's got the tables in the back of it that we put together many years ago, which George will go into details on some of that briefly. Post-TMI, and it's been a well established document that's been used for years. So, but with the advent of advanced reactors coming onboard, you know basically this document, Rev. 3 is designed for light water reactors. These new advanced reactor designs are other than light water reactors, in some cases. So we need -- so the industry felt as though a little broader guidance was needed.

And so we have attempted to endorse that with this standard. We considered several options and approaches to it because there's some things that are a little unusual about it which George will talk about in detail. What we'd like to do is just to capitalize on this opportunity to share this with you and get your reaction to it. Just it would make us feel I guess a little more comfortable. The next process is to send it out for industry comments. So NRR has

reviewed it and commented on it. OGC has also reviewed it and commented on it, but as I say, we haven't sent it out yet, so it's not quite gelled yet.

CHAIRMAN APOSTOLAKIS: You mean for public comments?

MR. KEMPER: For public comments, yes. Okay?

CHAIRMAN APOSTOLAKIS: Yes, great.

MR. KEMPER: So with that, George? Go ahead and get started.

MR. TARTAL: My name is George Tartal. I work for the Instrumentation and Control Section of the Office of Research. I've been with NRC for about a year, and before coming to NRC I had 13 years of experience in design engineering in the private sector.

CHAIRMAN APOSTOLAKIS: And you still want to stay with the NRC after a year?

(Laughter)

MR. TARTAL: I'm sorry.

CHAIRMAN APOSTOLAKIS: That's okay. It was a good decision joining the agency after? We are allowed to joke. Makes long sessions easier to take.

MR. TARTAL: Can you hear me better now? So as Bill mentioned, the reason we're presenting this

guide is because we're seeking the committee's verbal interaction on the approach taken in the content of the draft guide. First we'll provide a brief background on the history of accident monitoring, then discuss the current revision, Rev. 3 of Reg Guide 1.97. Then we'll provide a brief overview of IEEE Standard 497-2002, which is a revised standard for the selection, performance, design, qualification, display, and quality assurance criteria for accident monitoring. Then we'll describe the draft guide presented for discussion today, Draft Guide DG-1128, focusing on the regulatory positions and the issues the staff addressed in trying to endorse the standard in the guide. I'll describe the approaches the staff considered for the draft guide, followed by a conclusion and a request for any additional comments or questions on the approach and content of the guide.

10 C.F.R. 50, Appendix A, Criteria 13, 19, and 64 require instrumentation be provided to monitor variables in systems under accident conditions. Reg Guide 1.97 was issued as the effective guide in August of 1977, and provided general design and qualification criteria for accident-monitoring instrumentation. The accident TMI II happened in 1979. Lessons learned from TMI II and post-TMI action plan, NUREG-0737,

later codified in 10 C.F.R. 50.34(f) resulted in Revision 2 to the Reg Guide 1.97 in December of 1980. Revision 2 was to be implemented via NUREG 0737. A later revision, Revision 3 then reorganized the design and qualification criteria into tabular format, and revised some radiation-monitoring variables. It was issued 22 years ago in May of 1983 and is still the current source of accident-monitoring criteria for nuclear power plants.

Rev. 3 endorses ANS Standard 4.5-1980, which has since been withdrawn as now an inactive standard. And I'd like to briefly review the variable types and categories in the current guide since we're going to talk about them in a later slide.

CHAIRMAN APOSTOLAKIS: So the last revision was in 1983?

MR. TARTAL: Yes, that's the current revision.

CHAIRMAN APOSTOLAKIS: That's an interesting situation, that we're endorsing a standard that is now inactive. What does that say?

MR. TARTAL: That was -- that's the current guide right now. We're not talking about the draft guide. The current guidance is Rev. 3. The draft guide is going to become Rev. 4.

CHAIRMAN APOSTOLAKIS: Yes. I understand that.

MR. TARTAL: So that the current guidance is 22 years old.

CHAIRMAN APOSTOLAKIS: No, but I'm saying in 1980 I guess, no in 1983 we endorsed an ANSI standard that has been withdrawn.

MR. KEMPER: That's correct.

MR. TARTAL: It's since been withdrawn, yes.

CHAIRMAN APOSTOLAKIS: Why? Was it wrong, or why was it withdrawn?

MR. TARTAL: It was withdrawn because Rev. 3 of the reg guide became the sole source for accident-monitoring criteria. It really wasn't needed. Rev. 3 was so prescriptive.

CHAIRMAN APOSTOLAKIS: Oh.

MR. KEMPER: It became the de facto industry standard.

CHAIRMAN APOSTOLAKIS: Okay.

MR. TARTAL: Accident-monitoring variables prescribed in Tables 2 and 3 of the guide are organized by variable type. Type A are for planned manual actions with no automatic control. They're plant-specific and an example would be reactor coolant

level for monitoring core cooling. Type B are for assessing plant critical safety functions. An example is RCS pressure for monitoring RCS integrity. Type C for indicating potential or actual breach of fission product barriers. An example is primary coolant radioactivity for monitoring fuel cladding integrity. Type D for indicating safety system performance and status. An example is high pressure injection flow. Type E are for monitoring radiation levels, releases, and environs, with an example being plant vent radiation for monitoring airborne releases.

The design qualification criteria applicable to each variable are determined by an assigned category. Category 1 is for indicating the accomplishment of a safety function, and analogous to safety-related instruments. Category 2 is for indicating safety system status, and analogous to augmented quality-related instruments. Category 3 for backup and diagnostic variables, and analogous to non-safety related instruments. So with this prescriptive list of variables to monitor, and comprehensive set of design and qualification criteria to be met, Rev. 3 has become the de facto standard for accident-monitoring criteria in the industry.

With digital instrumentation being more

frequently employed in nuclear power applications, and with the new and advanced plant designs being considered for future licensing, a more flexible approach to accident-monitoring was desired by the industry. IEEE Standard 497-2002 was created to consolidate the criteria from inactive Standards ANS 4.5 and IEEE Standard 497-1981, as well as from Reg Guide 1.97 Rev. 3, and to update the criteria to the current state of technology. It provides a technology-neutral approach intended for advanced design plants. It takes a performance-based non-prescriptive approach to the selection of accident-monitoring variables. The prescriptive tables of BWR and PWR variables have been now replaced by variable selection based on design basis accident mitigation functions. This is the most significant change from Rev. 3. The selected variable type then determines the applicable performance, design, qualification, display, and quality assurance criteria. The standard reference is other recent industry standards in the criteria, and also provides criteria for the use of digital instrumentation. And the next slide provides a brief overview of this criteria.

The definitions for variable types A through E are similar to the definitions in Rev. 3 of

the guide. Some typical source documents are also referenced for each variable type, like EOPs, EPGs, AOPs, etcetera. Performance criteria in the standard include range, accuracy, response time, duration, and reliability. Design criteria include single and common cause failure, independence, separation, isolation, power supply, calibration, and portable instrumentation. Qualification criteria include environmental and seismic qualification for fixed and portable instruments. Display criteria include display characteristics, identification, display types, and recording. Finally, quality assurance criteria are given. The significant differences here in the criteria from that of Rev. 3 are new criteria for selection, additional criteria for single- and common-cause failure, guidance for use of portable instruments, and examples of monitoring channel displays.

This Draft Guide DG-1128 is the proposed Rev. 4 of Reg Guide 1.97. It was prepared as a response to a user need request from NRR. RES and NRR have worked together to come up with an approach that can be effectively implemented and regulated for new and current plants. The draft guide endorses IEEE Standard 497-2002 with exceptions and clarifications.

It's intended for new nuclear power plants with conversion to this new method by current operating plants on a comprehensive and strictly voluntary basis. And we'll talk about that in a minute. Next we'll discuss the five regulatory positions against the IEEE standard.

The first regulatory position addresses the question `How might current operating plants using Rev. 2 or 3 of the Reg Guide 1.97, how might they apply the criteria in IEEE 497?' The standard states it's intended for new plants, but, quote, "The guidance provided in this standard may prove useful for operating nuclear power stations desiring to perform design modifications or design basis modifications." The staff thinks that current plants may be interested to see if and how they can use the new guidance. The problem is the standard doesn't tell you how the current plants might use it. It tells them they can use some of the guidance. But what if current plants wanted to use all the guidance and convert to the new method? By "convert" what we mean here is moving from the current licensing commitments in Rev. 2 or 3 of Reg Guide 1.97 and revising their accident-monitoring program to the criteria contained in Rev. 4. The standard, since

it's intended for new plants, does not provide any guidance in translating from variable types and categories as they have in Rev. 3 to only requiring variable types in the IEEE standard. Since the categories do not directly correlate to variable types, the staff compared the variable types in associated categories, and concluded that generally Types A, B, and C are Category 1, Type D is Category 2, and Type E is Category 3. But there are some exceptions to this translation. The example shown here is PWR Subcooling Margin Monitor. It's a Type B Category 2 variable. If they were to convert this variable, would it become a Type B, or a Type D, or something else? The variable selection process would have to make this determination on a case-by-case basis. Furthermore, even if the variable type doesn't change, the individual criteria for that particular variable type may be different, and the converted variable would need to meet all the applicable criteria in the standard for that variable type. For current plants to convert some of the individual variables may require physical modifications as well as licensing basis changes. The new criteria may be more or less stringent than the current criteria, depending on the new selected variable type and the

existing variable's assigned category. As a result, we cannot intend this guide for current plants, but current plants may convert on a voluntary basis. The staff also feels that partial conversions of one variable or system could result in the potential for some variable or system interactions to be left un-analyzed and un-monitored, and hence conversion should be comprehensive of the entire accident-monitoring program. As a result, the draft guide states it's intended for new plants, and conversion for current plants may be done on a comprehensive and strictly voluntary basis by the licensee.

MEMBER BONACA: Yes. I mean, as I review this part, I still get confused about how you go from one to the other.

MR. TARTAL: It is confusing. It's not straightforward.

MR. KEMPER: Yes. And to add more to the confusion, you know, this is a new process. It really hasn't been worked out yet, right? So there's no plants out there with Rev. 4?

MEMBER BONACA: The most confusing thing was, I mean, so many of the changes in 1983 were tied to the issues that came out of TMI.

MR. TARTAL: Yes.

MEMBER BONACA: And you know, I was trying to trace back on how you would deal still with those issues in an explicit fashion based on this new guidance, and our regulatory position, you can trace it easily. This doesn't seem to be specific requirement pointing into that direction, while the old reg guide clearly had pointers there. You could see why they did certain things because of the experience of TMI. So it's a little confusing. Do you expect that the people with current plants would go this new approach?

MR. KEMPER: Yes. I've received a couple of calls so far from the BWR owners group representatives. And from indications I've gotten through those calls that they're waiting for this to be issued so they can evaluate, I guess, what they want to do, if anything, to the current generation plants.

The other point here too is by having a situation where plants are straddled, if you will, part of their post monitoring PAMI instrumentation is in Rev. 3, complies with Rev. 3, and part of it goes to Rev. 4. It'd be very difficult I guess from an inspector's standpoint to go out and actually audit, you know, what the licensing criteria is. And

secondly, NRR would really have quite a burden placed on them for these submittals as they come in, you know, one by one, or two instruments here, one instrument there, trying to comply with this new standard and leave the rest of them where they are. So that --

MEMBER BONACA: That's another issue that I was thinking of, you know. Again, this piecemeal application, if it happens, takes existing plants away from some level of standardization that we have been able to implement in these plants to whatever degree we could. And that standardization I believe is responsible for improvements in safety performance, just because there is a lot of news of lessons learned from sister plants. And this could be radically different. I mean, you could see departures that would take somebody pretty much away from the experience. Anyway, it's just an observation.

MR. TARTAL: So the second regulatory position the staff addressed was the IEEE Standard's requirement for maintaining channel calibration during an accident. The standard requires maintaining instrument calibration by means of re-calibration, proper calibration interval specification, selecting equipment that does not require calibration, or by

cross-calibration with other channels having known relationship to that variable. The staff believes that although conceptually a good idea, plants should not be required to maintain calibration during the accident. Instead, the draft guide states that the plants should design accident-monitoring channels to the extent possible with the ability to maintain calibration during an accident.

The third regulatory position addresses IEEE Standard's future work section on severe accidents, and how it relates to selection criteria. The standard does, however, include the requirement for Type C variables to have extended ranges, which was a post-TMI action item now in 10 C.F.R. 50.34(f). The agency's severe accident policy does not require mitigation of severe accidents, and hence there are no requirements to monitor severe accidents. However, the draft guide incorporates the language from NUREG-0660, which is the post-TMI action plan, into the criteria to clarify the requirement for extended ranges for Type C variables, but does not further address severe accidents.

The fourth regulatory position addresses the IEEE Standard's exclusion of contingency actions from the variable selection process. Contingency

actions are most commonly associated with those additional actions and EOPs used when primary success paths have not been successful. The IEEE standard assumes that all contingency actions are to mitigate action conditions that are beyond the licensing basis of the plant. But the staff doesn't want to unnecessarily exclude contingency actions from the potential list of variables to monitor if some of those actions could be a potential accident-monitoring variable in accordance with the given criteria. Therefore, the staff feels that this restriction toward contingency actions should not be endorsed. Instead, the licensee should consider all EOP actions for design basis events during the variable selection process, allow the selection criteria to determine if the variables used for the contingency action can be excluded.

The fifth regulatory position is a carryover from Rev. 3 of Reg Guide 1.97, and addresses the number of points of measurement for a variable. The IEEE standard does not address a number of points of measurement for a variable like Rev. 3 did. The regulatory position states that the number of points of measurement for each variable should be sufficient to adequately indicate the variable value. In other

words, for example, if you were to measure containment temperature due to the size of containment space, you wouldn't want to use a single point measurement and say that's representative of everywhere inside containment. You'd want several measurements at various locations.

Next I'll briefly describe the four approaches the staff considered to solving this need for a more flexible source of accident-monitoring criteria. One approach was to take no action. Reg Guide 1.97 would remain at Rev. 3 for current and new plants, and IEEE 497 would not be endorsed. That solution may be adequate for the fleet of current operating plants, but the prescriptive variable list and outdated criteria of Rev. 3 wouldn't be of much use for a licensee of an advanced design plant. So the staff did not choose this approach.

The second approach the staff considered was to revise Reg Guide 1.97 to incorporate all previously approved deviations which were generic to that particular design, as well as other clarifications and role changes as a means of updating the guide for current plants, and at the same time endorse IEEE 497 for both current and new plants. First, all the changes that I mentioned a second ago

have already been approved, so it would be unnecessary to go through the process of re-approving them in this guide. Second, providing separate guidance for current and new plants within the same reg guide could result in a confusing and ambiguous guide. Therefore, the staff did not choose this approach.

The third approach the staff considered was to have two reg guides addressing accident-monitoring. A new reg guide, 1.xxx endorsing IEEE 497 would provide accident-monitoring criteria for new plants, and Reg Guide 1.97 Rev. 3 would remain the reg guide for accident-monitoring for current plants. The first problem is the nuclear industry knows Reg Guide 1.97 is the sole source for accident-monitoring criteria. The staff feels that issuing a second reg guide also providing accident-monitoring criteria would be confusing to licensees and regulators. Second and more importantly, there are a number of regulatory documents which refer to Reg Guide 1.97 for accident-monitoring criteria, like 10 C.F.R. 50.49 and Reg Guide 1.89. And the staff would need to revise all the regulatory documents that refer to the Reg Guide 1.97 to also refer to this new Reg Guide 1.xxx. So the staff didn't choose that approach either.

The fourth approach the staff considered

was to revise Reg Guide 1.97 to endorse IEEE 497 for new plants as the standard says it's intended. Current plants would continue to use the guidance in Rev. 2 or Rev. 3 of Reg Guide 1.97, or voluntarily and comprehensively convert to the criteria in Rev. 4. The benefits of this solution are that it endorses the updated consensus standard for new plants, which Approach 1 didn't do; it would create clear and unambiguous guidance for new and current plants, which Approach 2 didn't do; and retain the industry-familiar name of Reg Guide 1.97 for new and current plants, which Approach 3 didn't do. As a result, this is the approach that the staff chose. Furthermore, NRR and OGC have reviewed the draft guide, and have no technical or legal objections to the content approach in the draft guide.

In conclusion, Draft Guide DG-1128, the proposed Revision 4 to Reg Guide 1.97 endorses the current industry standard IEEE Standard 497-2002 with exceptions and clarifications. It's consistent with and provides a method for meeting the NRC's requirements. Standard Review Plan Chapter 7 will require updating for the new revision of the guide. The revision is intended for new nuclear power plants, and any current plant wishing to convert to the

criteria contained within may do so on a comprehensive and voluntary basis. There are no backfit issues associated with this revision. And finally, we ask the subcommittee if there are any additional comments or questions that you have before we proceed with issuing the draft guide for public comment.

MEMBER WHITE: Excuse me, could you clarify what you mean by "no backfit issues"?

MR. TARTAL: Since the draft guide is intended for new plants, it doesn't affect the current plants. Backfit issues are associated with current operating plants.

MR. KEMPER: We've tried to emphasize voluntary use for current generation plants as the only way that we would -- the way we are endorsing the standard. To be very clear about that.

CHAIRMAN APOSTOLAKIS: Any comments? No?

MEMBER BONACA: I just have a question. I mean, you know, I can see how the licensee could take this new approach, okay, through some way that wasn't clear to me how it was easy it's going to be. He would then choose certain issues of the protection system or ESF and so on and so forth features. Do you envision that there was a transition of that type by many at some point the NRC would feel compelled to go

back to a more prescriptive approach for individual types to plant? I'm talking about the type of plant, like you know for example for PWRs, they finally decided that everyone had to have this specific measurement. Everybody had to have the same.

MR. TARTAL: Those tables were initially put into the reg guide because the industry didn't understand how to implement the previous revision of the reg guide.

MEMBER BONACA: Okay.

MR. TARTAL: It gave general design and qualification criteria, and at that point accident-monitoring was still in its infancy. People didn't understand how to use the general criteria. So to make it more clear, the NRC came out with Rev. 2 which had the prescriptive list of variables.

MR. KEMPER: And I think that history has shown -- Barry you can speak up here if you'd like -- that as time has gone on, there's been many exceptions requested and granted to the prescriptive list in Reg Guide 1.97.

MR. TARTAL: Deviations.

MR. KEMPER: Yes, deviations by various NSS-type or plant-specific issues and so forth. So this new performance-based criteria hopefully will

eliminate that. Each plant will do their own analysis unique unto itself, and then of course NRR will have the task of reviewing and approving that.

MEMBER BONACA: And I agree that the understanding of plant behavior has changed significantly, so that will be acceptable. Thank you.

MR. TARTAL: Okay. Other questions?

CHAIRMAN APOSTOLAKIS: Okay. Thank you very much.

MR. KEMPER: Thank you.

CHAIRMAN APOSTOLAKIS: Sam, maybe you can help us here. Can we start the next -- the afternoon session a little earlier?

MR. DURAISWAMY: No.

(Laughter)

MEMBER WHITE: Does that mean you need more dialogue?

CHAIRMAN APOSTOLAKIS: Okay. We'll recess then until 12:30.

(Whereupon, the foregoing matter went off the record at 11:08 a.m. and went back on the record at 12:30 p.m.).

CHAIRMAN APOSTOLAKIS: We're back in session. The next item on the agenda is a short presentation by Mr. Kemper on software quality

assurance. Correct?

MR. KEMPER: That's correct.

CHAIRMAN APOSTOLAKIS: Okay.

MR. KEMPER: Thank you. Well again, I'm Bill Kemper, the section chief for the Instrumentation and Control Engineering Section of Research. And since we've got some new members here, I'll just give you a quick background of myself. I've been with the agency for just a couple of years. I'm a relative newcomer. I spent 29 years in the nuclear industry before that, worked at three different utilities, and three different power plants, and spent a lot of time, done a lot of things in my career, but a lot of it was in operations and instrumentation and control engineering. So it's a pleasure for me to be here working with this agency on the regulatory side of the business.

So at any rate, I only have 15 minutes to speak, so I will try to get through this on time. I just wanted to provide a brief discussion, kind of an overview of what we're trying to accomplish here in this area of software quality assurance. The diagram you see before you is out of the research plan. This covers the activities that are currently scoped out for Section 3.2 of the research plan. Right now we

have three initiatives that we're focusing on in this area. You'll receive detailed presentations following mine on each one of these areas. So at any rate, there's more time for more questions as they come up in each one of these areas.

And listening to the presentations this morning, actually I kind of -- I'd like to build on some of the statements that were made earlier about the research programs. What we tried to do is put this presentation together such that we can explain what the agency is doing now, what the areas for improvements might be, and then what we intend to do about it, it boils down to, okay? So.

CHAIRMAN APOSTOLAKIS: This will be a good template for all the presentations.

MR. KEMPER: So to provide some -- I'm sorry.

CHAIRMAN APOSTOLAKIS: Oh go ahead.

MR. KEMPER: Yes, to provide some background on the current process for evaluating software quality of licensee applications, the NRC SRP Chapter 7, Standard Review Plan, Revision 4 which was issued in June of 1997 provides the regulatory framework for the review and approval of digital safety systems. As part of its review of digital

safety systems, NRC evaluates safety-related software quality by reviewing the developmental process, for example verification and validation testing, configuration management programs, and software development products, such as software requirement specs, software design documentation, test plans, requirement traceability matrices, those sort of things. In other words, the agency reviews the software developmental processes and products produced by the vendors and the licensees themselves. Now, I think we're all in agreement, the SRP is adequate to provide guidance, in other words, what to review, to the staff in performing safety reviews that pertain to digital safety systems.

The review and approval of digital systems currently depends on qualitative evaluations of digital system features and development processes. Software quality assurance evaluations are performed manually, without the aid of assessment tools or other means of obtaining quantitative measures of software quality. And also, the SRP Chapter 7 Branch Technical Position 14 identifies digital system development attributes that should be reviewed, but does not really provide detailed guidance on the process for confirming that the software conforms to the

acceptance criteria.

CHAIRMAN APOSTOLAKIS: So this slide then is what we're doing now, and what we need to do.

MR. KEMPER: This is the delta, if you will.

CHAIRMAN APOSTOLAKIS: Good.

MR. KEMPER: This is, as we see it, the area for improvements that we're trying to set the foundation for that. So as I've stated, the SRP is a very thorough document, very thorough compilation of what requirements must be satisfied. What we're attempting to conduct research on is to provide the reviewer with information about how the criteria should be satisfied, and also how much is good enough, quite honestly. As Mike Waterman said earlier in his presentation, a lot of the reviews is a function of what the reviewer has within himself or herself in terms of meeting these criteria.

CHAIRMAN APOSTOLAKIS: Now, when you say in the second bullet software quality assurance evaluations are performed manually, you envision in the future the reviewer to have computer help?

MR. KEMPER: That's true. I'm going to get into that very shortly here. In the next slide or two. So NRC reviews the results of software

development processes and safety assessments, but the reviews do not include a means for independent assessments of software quality assurance. Now, by "independent" I mean what we're trying to provide is a method for reviewing software that does not just rely on licensee- or vendor-produced products. We hope to provide tolls that will provide another dimension to the agency's capabilities to review software. For example, when the licensee submits a new fuel design for review, the agency not only reviews the code and documentation that the licensee used for the new fuel design, but the NRC has its own codes that it can run independently to verify what the licensee has concluded. And you can make the same statement in the PRA business. The agency has its own PRAs to use to validate licensee activities pertaining to risk. We don't have tools like that in the I&C business, so that's what we're proposing to do is try to create some of those tools for independent assessments.

So given the complexity and sophistication of current digital safety systems, the goal of this research program is to provide independent assessment methods and objective acceptance criteria that can supplement and augment the existing guidance in

Chapter 7 of the SRP. These words, you've heard these several times. We're going to continue to make this statement as we go through our projects.

So this information can be provided as formal review procedures for verifying consistency with the SRP guidelines, which could also improve effectiveness and consistency of software quality assurance evaluations and reviews.

MR. ARNDT: Let me jump in here for a second. The point here is that if we have these extra tools, or additional methodologies, or additional information, we don't have to use them in every case. But where we want additional information, or where it would be useful, or there's a particular issue, the idea is to have these available so that we can do additional work if we feel that's justified.

MR. KEMPER: Okay. Also, the current state-of-the-art in software system safety assessment includes a number of methods and tools for quantitatively assessing the quality of software. For example, there are software system analysis techniques such as Petri-net analysis, Markov analysis, dynamic flow modeling, being used in software modeling techniques right now. Tools such as software metrics, formal verification methods, and testing techniques,

such as data flow testing, fault injection, and mutation testing, are being used for software design analysis techniques to ensure that the software system works in a particular way.

So what we're trying to accomplish is to review what software quality assurance methods and tools are out there being used in other sectors of the process control industry. And we will then, if possible, adapt these tools for deployment on software systems within the nuclear industry.

CHAIRMAN APOSTOLAKIS: I wouldn't use the word "quantitatively" on your first line. There are a number of -- like, I don't think formal verification methods are quantitative. I mean, they're logic.

MR. ARNDT: They're logic systems to verify that --

CHAIRMAN APOSTOLAKIS: Quantitative means you produce numbers. So I mean, you can still make your point by deleting the word "quantitative".

MR. ARNDT: We can do that.

MR. KEMPER: I guess the point here though is it's a process. It's a consistent process.

CHAIRMAN APOSTOLAKIS: I understand.

MR. KEMPER: It's an algorithm, right? In other words, it's a methodology that's --

CHAIRMAN APOSTOLAKIS: These are methods that --

MR. KEMPER: -- that's not the function of the individual, the human being.

CHAIRMAN APOSTOLAKIS: Structured methods.

MR. KEMPER: Structured, exactly, very good.

MEMBER GUARRO: Structural, formal.

MR. KEMPER: Exactly.

CHAIRMAN APOSTOLAKIS: All of them are formal. Right? Even the third bullet there. Because you insert the word "formal". Software metrics, I don't know what you mean by that.

MR. KEMPER: We're going to explain that to you in just a minute.

CHAIRMAN APOSTOLAKIS: Good.

MR. KEMPER: Okay. So therefore, research in this area will focus on assessing possible analysis methods that are currently used in design and analysis of safety-critical software systems to use in the regulatory process. We intend to focus on methods that have likely short-term application without the need to do extensive development and apply these to nuclear industry applications. For example, fault injection testing has been used by a number of

industries, including some nuclear platform suppliers. Formal methods have been used in several industries to support safety-critical applications. Software metrics are currently being used for software quality control and continuous improvement activities in organizations that have programs that are capability maturity model level 4 and 5 respectively. In fact, all military vendors right now are required to have a CMM level 3 program in order to even bid on a contract. So we're just trying to build on these tools and technologies that are out there. And also, any nuclear supplier and vendor should be at least a CMM 3 level because they have a well-defined program per 10 C.F.R. 50 Appendix B, and so they should be ready and capable to implement metrics.

And in summary, this research area currently focuses on three initiatives to develop independent methods of assessing software quality and/or reliability: the use of software metrics to evaluate quality, the use of fault injection techniques to evaluate digital system dependability, and to provide technical guidance and review procedures for evaluating self-testing features in digital systems. Now, self-testing features is not really an independent testing method in and of itself.

This is really a review criteria issue. So what we want to do is investigate the self-testing methods and technologies that are being used right now in the industry, and try to get a better idea of what are the best testing schemes that we're aware of, and how much reliability is gained from the various self-testing schemes, considering the failure probability presented to the software system due to the added complexity associated with the self-testing software itself. In other words, how much benefit is gained for the extra complexity. Right now we don't have any information to build on in that arena.

CHAIRMAN APOSTOLAKIS: So why did you decide not to pursue formal verification methods? That's the only one you're leaving out isn't it?

MR. ARNDT: Well, we're choosing to look at particular aspects of particular projects. We looked at formal methods through our cooperative agreement with Halden because that's part of their research program. The results to date didn't appear to be as promising as other methodologies. We continue to keep track of formal methods through our cooperative agreement through Halden. To my knowledge, I'm more than happy to be informed, there was a lot of work in this area in the `80s and `90s,

some successes, then it kind of fell out of favor for awhile. It's starting to become more popular now that some of the tools are much more sophisticated. So as with anything else in the research program plan, as we have the resources, we're going to look at whether or not any particular methodologies may be useful. If they do appear to be useful, then it will get rolled into the next upgrade a year from now, or two years from now, whenever.

CHAIRMAN APOSTOLAKIS: Is the work that you're doing with Halden mentioned in the plan? I can't remember.

MR. KEMPER: Yes, I think it is mentioned in the plan.

MR. ARNDT: It's part of, I think, the cooperative international agreements, which is in Section 3? Probably 3.7.

MR. KEMPER: Although there's no specific projects that are the outcome of that directly in and of themselves. We use that right now as supporting information for background and to integrate into other existing projects. But I think Steve's making a good point here. The idea of this research plan is it's a flexible document. So if we have good reason to believe that formal methods is an area that we should

pursue, then we will certainly do that.

CHAIRMAN APOSTOLAKIS: Well, I mean it depends on what you call formal methods. Because of course, most people would think of the logic --

MR. ARNDT: Proof calculuses and things like that.

CHAIRMAN APOSTOLAKIS: Find errors and so on. Or confirm that things are self-consistent. But I recall that the Canadians adapted these methods. They didn't quite use formal methods to prove correctness, but they borrowed heavily, you know, developing tables and all that.

MR. ARNDT: Yes. They use it as a design criteria, basically.

CHAIRMAN APOSTOLAKIS: Yes. I mean, are you familiar with what they have done?

MR. ARNDT: Yes.

CHAIRMAN APOSTOLAKIS: Is there anything useful there?

MR. ARNDT: I've read some of the work. Also, the Brits did some work in that area on Sizewell as more of a design methodology as opposed to a formal correctness proof.

CHAIRMAN APOSTOLAKIS: Okay. All right. You done?

MR. KEMPER: Okay. Almost. So at any rate, to conclude, these research projects will provide objective acceptance criteria and review procedures that augment and supplement existing SRP guidance for approving or denying digital safety system license applications. And that's the hardest part. When we deny something, we need to have a solid foundation to build on. So that really concludes my short overview of this area. If there's --

CHAIRMAN APOSTOLAKIS: So what is the distinction between quality assurance and the risk part of it?

MR. ARNDT: The big issue is quality assurance is the effort to assure or get a level of confidence that the software is performing safety functions appropriately.

CHAIRMAN APOSTOLAKIS: Without quantitative estimates.

MR. ARNDT: Without necessarily having quantitative estimates. That doesn't mean you can't have quantitative estimates, it's just not the primary objective of quality assurance.

CHAIRMAN APOSTOLAKIS: Well, let's say that you find yourself sometime in the future, you really trust the risk methods. Then all this would go

away, wouldn't it?

MR. ARNDT: No.

CHAIRMAN APOSTOLAKIS: If I trust a method, and the method tells me I have a 10^{-5} or -6 variability, I have a high level of confidence that this is pretty good.

MR. KEMPER: Well, but the quality I think is an underlying principle that has to be preserved for those risk performance measures to be valid. Okay? The failure probably is predicated on certain underlying notions.

CHAIRMAN APOSTOLAKIS: If it were not preserved, would I get a number as low as 10^{-5} ?

MR. ARNDT: Presumably not --

CHAIRMAN APOSTOLAKIS: No.

MR. ARNDT: But the point is we're not a risk-based organization, nor are we likely to be.

CHAIRMAN APOSTOLAKIS: I put you in a hypothetical situation.

MR. ARNDT: Okay.

CHAIRMAN APOSTOLAKIS: So this it seems to me is important because we cannot do the other thing. We cannot really estimate risks with any kind of confidence.

MR. ARNDT: Well, you get into the same

basic state that we have in any part of the business. I mean, we have Appendix B, and we don't -- we can grade quality, if you will, by risk, but you don't get rid of quality assurance.

CHAIRMAN APOSTOLAKIS: No, you don't.

MR. ARNDT: Because you need to have that understanding that the process is working, that there was appropriate --

CHAIRMAN APOSTOLAKIS: Because a lot of these things cannot be modeled in the PRA.

MR. ARNDT: That's right. And even if they can be, you're never going to have 100 percent confidence. So there's several different ways you attack the problem. The purpose of this program is simply to use the software engineering methods that are out there to try and make software quality assurance evaluations better.

CHAIRMAN APOSTOLAKIS: But the fault injection technique, for example, it has, you know, you inject the faults and see what happens and so on. And then they go on to do some numerical calculations. You don't mean that the whole package here, I mean, part of it may be useful, part of it may not.

MR. ARNDT: Yes. The real issue in these programs -- and I don't want to talk through all the

programs because we have presentations for that -- is to gain a better understanding of both the process and the product itself, understand how the system works or doesn't.

CHAIRMAN APOSTOLAKIS: Okay. So you have presentations on each one of these?

MR. ARNDT: Yes.

CHAIRMAN APOSTOLAKIS: Okay, great. Let's go on then.

MR. KEMPER: Okay. As a matter of fact, the next presentation is by Norbert Carte, and Steve Arndt also will participate in this, and also this is Ming Li from the University of Maryland.

MR. CARTE: Hello. My name is Norbert Carte. I am also in the I&C section, Engineering section of the Engineering Research Applications Branch. I've been with the NRC since early February, and prior to that I spent 13-plus years performing verification and validation of various digital systems in the nuclear industry. I'll be presenting today with Ming Li, one of the researchers from the University of Maryland. And I'll allow him to introduce himself.

DR. LI: My name is Ming Li. I'm a research associate at the Center for Reliability

Engineering at the University of Maryland in College Park. I've been a key researcher for this project since 1998. I hold a Bachelor's degree in Electrical Engineering, and the Master's in Systems Engineering, and the Ph.D. in Reliability Engineering. My research interests include software engineering, reliability engineering, software measurement, software testing, and the PRA. Thank you.

CHAIRMAN APOSTOLAKIS: So I take it you will talk about the metrics?

DR. LI: Right.

MR. CARTE: Ming will be talking about two metrics in detail, and I'll be giving an overview of the program itself. So we'll start off with a discussion of the issues facing the NRC, some of which you've heard previously, as well as the basis of the current engineering project, and then discuss two metrics in detail, and follow on with a brief discussion of future work and conclusions.

The basic issue facing the NRC is regarding the increasing size and complexity of submittals. And this will result in an increased workload, and with the limited staff that could present some problems. Software is currently being used in more systems as well as an increase in the use

of self-checking software and other techniques result in more complex systems. Also, with the use of commercial off-the-shelf equipment we have more powerful development environments, and that means that software programming is becoming more complex, or abstract, as well as many of the details are becoming hidden. Software engineering methods are also becoming more powerful and usable, and therefore can be used to address these issues.

CHAIRMAN APOSTOLAKIS: Now, are these comments true for existing reactors? I mean, are we really using complex software? Not for future reactors. I am talking about, you know, control and all that, feedback. I mean, what is the level of the sophistication of the software that are being used in safety-related functions these days?

MR. CARTE: Well, the question is not necessarily just what is currently being used, although I believe there are some 30 systems that have been approved. There are, in general, three SERs, Triconex, Westinghouse, and Teleperm TXS which propose using development environments and systems, and the potential application is for plant-wide modernizations. And the obsolescence issue will result, possibly, in many plants wanting to do

complete plant-wide modernizations. And there are some -- but currently what we see are system-level modernizations.

CHAIRMAN APOSTOLAKIS: Are there any plants right now that are using digital software in safety-related functions?

MR. CARTE: Safety-related. I think Vogtle has a diesel sequencer that uses a Westinghouse ABB Advant system.

MR. KEMPER: Sure, the CE System 80 Plus design. It's got a compression calculator. Let's see. What is it, the Eagle?

CHAIRMAN APOSTOLAKIS: What is that system?

MR. KEMPER: Eagle 21.

MR. WATERMAN: Eagle 21 is a reactor protection system.

MR. KEMPER: Yes. There are numerous spotted applications out there, but it's not on a generic-wide basis.

CHAIRMAN APOSTOLAKIS: So the reactor protection system is basically monitoring and then SCRAMming?

MR. KEMPER: Right, it's a trip system. Exactly. But like the core protection calculator is

--

CHAIRMAN APOSTOLAKIS: But aren't these relatively simple systems?

MR. KEMPER: Well, the basic function of tripping, you know, comparing a set point to a parameter and then tripping your relay is, but like the core protection calculator, it's got a fair amount of sophistication involved with calculating that variable trip set point.

MR. WATERMAN: And those have always been digital in several plants.

MR. KEMPER: The point here though I think that Norbert's trying to make, and excuse me for breaking in on you here Norbert, is that increasing complexity and size of submittals. There's nothing to prevent licensees from making submittals for plant-wide upgrades. In fact, when I was at Calvert, that's one of the last projects that we concluded was a plant-wide digital upgrade project for, you know, cost us $60 million over the next 10 years. So this is what's going on out there in the industry, and that's what we're being subjected to. Those submittals could come at any time.

CHAIRMAN APOSTOLAKIS: Is the Oconee license amendment request that was mentioned this

morning the first serious step towards using digital I&C in safety systems?

MR. ARNDT: Well, I don't know if you'd call it the first serious step, but it is a very large step that will include RPS and SFAS and other systems.

MR. KEMPER: I believe that's true though. That's a good way to quantify it. I mean, others, I think Callaway approached this once, and then they withdrew after some interaction with the staff.

CHAIRMAN APOSTOLAKIS: Because the regulatory stuff more seems to feel that this is really --

MR. KEMPER: Yes, I think it is.

MR. CARTE: Well, it also represents a change. The fact that you're integrating two systems into one system. You're integrating the RPS and the SFAS. And digital systems allow for that sort of thing.

CHAIRMAN APOSTOLAKIS: Yes. Yes, I agree. I'm trying to get a picture. Anyway, keep going.

MR. CARTE: Okay. So as has been gone in a little more detail this morning, the current review process is basically a software development review process as well as some sample threat audits that are selected by the reviewer. Standard review plan is a

generic plan, and it requires an application-specific review plan. The reason I point that out is there are different programming paradigms, such as structured programming, for instance something programmed in C, object-oriented programming in C++, and programmable logic controllers. Each represent a different paradigm, will have different vulnerabilities or weaknesses and different strengths. And therefore it might be better to have specific review criteria for different paradigms, as well as potentially measures.

The reg guides that currently endorse generic IEEE standards, in other words they're not programming paradigm-specific, as well as the current standard review plan does not address the use of measures.

CHAIRMAN APOSTOLAKIS: I noticed both in the previous presentation and this one, you guys are very careful to point out, you know, this is where we are, this is where we're going. I didn't get that impression from the plan that I reviewed. Is the new version going to be as explicit? I understand you are revising it now, right?

MR. KEMPER: Yes, we are. And --

CHAIRMAN APOSTOLAKIS: Because this is really the way it ought to be. This particular issue,

this is what's happening now, these are the issues, and this is how we're going to help. I sense there is a disconnect there.

MR. KEMPER: Well, we --

CHAIRMAN APOSTOLAKIS: But I reviewed the earlier version I must say, so I know that you are revising it now. But it would be nice to spend a few extra hours, Bill, to make sure that it's very clearly stated in each section where we are and where we're going. I think that's the main idea behind a good plan.

MR. KEMPER: I think that's absolutely right. We attempted to do that in the initial draft. We provided a background for each one of them which really addressed the issues, here's the problem statement, if you will, and then the task that we intended to accomplish. So certainly it's obvious we need to embellish that. We'll do that.

CHAIRMAN APOSTOLAKIS: That's all. Yes. Okay, let's move on. Boy, you're really slow, aren't you? You've been here only since February you say?

MR. CARTE: Yes.

CHAIRMAN APOSTOLAKIS: Well, we joke every now and then.

MR. CARTE: Yes. So the current research

goals. The objective of this research is to perform a large-scale validation of measures identified previously through previous research to quantitatively assess the quality of software.

CHAIRMAN APOSTOLAKIS: You know, this now raises the expectations. You say quantitatively. I'm looking for numbers.

MR. CARTE: Yes.

CHAIRMAN APOSTOLAKIS: Do you want to delete that word now, or?

MR. CARTE: No.

CHAIRMAN APOSTOLAKIS: Shall we keep looking for numbers?

MR. CARTE: Well, numbers in themselves aren't bad.

CHAIRMAN APOSTOLAKIS: Well, that's what quantitative means.

MR. CARTE: Yes. The question is how you use those numbers.

CHAIRMAN APOSTOLAKIS: No, no, no. I would like to know whether you produce them first.

MR. CARTE: That is the intent, yes.

CHAIRMAN APOSTOLAKIS: So this is all quantitative?

MR. CARTE: Yes.

CHAIRMAN APOSTOLAKIS: All right. Let's see. Okay.

MR. CARTE: That is, we envision the incorporation of measures to produce standardized quantifiable evaluations. Now, the question of what you do with those numbers relates to the acceptance criteria. How do you establish an acceptance criteria once you have a repeatable number generation system. And there are different ways of establishing acceptance criteria. Some are theoretical, and others include benchmarking it, or some combination of theoretical and benchmarking.

The purpose of this research is to be flexible as well, to look at measures that could be used by the licensee, the NRC, or both. And also, we want to address how you compare or combine different assessments. So when you look at a software design description, or a software requirements specifications, and have a quality determination, how do you compare those? Are you comparing apples and oranges? Or how do you compare the thoroughness or completeness of testing to the quality of the software requirement specification? One method of performing such a comparison is a Bayesian method, which basically relies on a probably or confidence, and then

using the Bayesian techniques to combine them. The other way would be to normalize the quality assessment in terms of some common measure or metric, such as defect density or reliability. The other goals of this research are to address the issues previously raised.

CHAIRMAN APOSTOLAKIS: You are not going to develop any methods that are usable by the licensees and not the NRC? I mean, you better rephrase that. You say they're licensee, NRC, and/or both.

MR. CARTE: Yes.

CHAIRMAN APOSTOLAKIS: Well, no. You're developing tools for the NRC, right? You are a member of this agency.

MR. CARTE: Yes.

CHAIRMAN APOSTOLAKIS: If the licensee wants to use them, fine. I can assure you that we'll --

MR. KEMPER: That's what we meant to say, actually.

CHAIRMAN APOSTOLAKIS: I know. I know. So change the words.

MR. CARTE: Okay. The use of metrics for quantifying software quality has a large basis in

literature. To give you an example, I've listed three IEEE standards regarding the use of measures and metrics. In particular I wanted to point out in 1061, one of the statements which says, "The use of software metrics does not eliminate the need for human judgment in software evaluations." So it is not the intent to replace human judgment, it's to provide more resolution, more information to the individual performing that judgment.

From that general literature and industry search, Lawrence Livermore Laboratory identified a pool of 78 measures. From that pool, the University of Maryland selected 30 measures, and categorized those measures in terms of the lifecycle phase to which they were applicable, as well as the semantic category, such as size and complexity. This was done in part to ensure all areas were covered, all lifecycles, and all semantic families.

They then elicited expert opinion in order to rank those measures and families. They also elicited peer review to evaluate the research performed. They also performed a preliminary evaluation which was published in the NUREG/CRITERIA that's identified, as well as wrote some publications in peer reviewed journals.

MEMBER WHITE: Excuse me. The peer reviewed journals, are those journals in which the software community normally publishes? So they're not just our industry journals, is that correct?

DR. LI: Yes.

MEMBER WHITE: Thank you.

CHAIRMAN APOSTOLAKIS: Like which one?

DR. LI: IEEE Transactions on Software Engineering.

CHAIRMAN APOSTOLAKIS: So you're going to tell us what it is, right? Soon.

MR. CARTE: Yes.

CHAIRMAN APOSTOLAKIS: Okay.

MR. CARTE: So the large-scale validation project being performed by the University of Maryland selected a sample of the measures. It is not validating all 30 measures. It selected that sample from the different classes of measures, some highly ranked measures, some medium, some low ranked measures, as well as different semantic -- from different semantic families. One example of a semantic family is the functional size, such as feature point, function point, or full function point, and complexity, such as cyclomatic complexity. And these measures were applied to all phases of the

software development lifecycle in a nuclear reactor protection system.

So the issues raised previously. In the NUREG itself, the issues raised identified during the peer review was that it was -- the preliminary validation was performed on a relatively software application. The application was not a nuclear safety system, which means that they looked at a low reliability system, as opposed to an ultra high reliability system. The benchmarking of the data did not use real operational profile, and it looked only at one phase of the software development lifecycle. And these issues are addressed in the current research project.

The ACRS addressed some of these issues, as well as some others. One is the ease of obtaining the metric. The current research will provide an evaluation of the ease of use for the metrics that they validated. A comment was software-centric versus a system-centric approach. We are more conscious or aware of the need to consider the entire system, and are looking at it from that perspective, although we are primarily looking at systematic failures that have a software origin.

Another issue raised was that the

uncertainty in the preliminary research was greater than the required reliability of the ultra high reliability systems. That is an issue we're conscious of, and we're looking at the research to address that, but some things to think about. With a low reliability system we had lower reliability numbers and higher -- and larger uncertainties than we would desire for an ultra high reliability system. The other issue is that this is not necessarily a new issue. If we have qualitative evaluations, there is always an uncertainty associated with a qualitative evaluation. The problem is we haven't specified what reliability is required, or we haven't talked about the uncertainty associated with that qualitative evaluation. So it's not necessarily a new issue, we're just trying to resolve that issue, and it becomes more visible when we start talking quantitatively. And I just want to point out that measures do not eliminate the need for human judgment.

The other ACRS comment was regarding the validity/robustness of the measures. So we are applying the measures to a different type of system, a different function, so we're looking at an RPS rather than a door entry system. We're looking at different programming languages, such as C & Assembler

versus C++. So with that I'll turn the discussion over to Ming.

DR. LI: Our current technical goal is to try to quantify software quality through software engineering measurement.

CHAIRMAN APOSTOLAKIS: You need a microphone if you're going to stand up.

DR. LI: I'll sit here, sorry.

CHAIRMAN APOSTOLAKIS: Yes, you keep talking and we will try to find him. Yes, you can use the cursor.

DR. LI: The philosophy behind this research is summarized as the answer to a question what determines software quality. In general, software quality is determined by the software product, the characteristics, in particular the defect remaining in the software, and how the software may be used. The way software is used is summarized using the concept of operational profile. Software product characteristics can be further determined by the product characteristics, for instance, what type of application is it, how big is the functional sizes. And the process characteristics, for instance, how good the developer's skills are, how tight the budget is, what development tools and methods are used,

etcetera. All these characteristics can be explicitly or implicitly described using software engineering measurements. Therefore, an obvious inference is software engineering measurements determine software quality.

CHAIRMAN APOSTOLAKIS: You seem to be focusing on numbers here.

DR. LI: Right, I'm going to talk about numbers shortly.

CHAIRMAN APOSTOLAKIS: But I don't care about the number of defects. If I have one that is fatal, that really bothers me. So it's not really the number. I mean, it's important to know the number, but --

DR. LI: Right, right --

CHAIRMAN APOSTOLAKIS: When do you focus on the significance of the defect?

DR. LI: Right, I'll talk about it shortly.

CHAIRMAN APOSTOLAKIS: You'll talk about it. Okay.

DR. LI: So the following steps are taken to pursue this technical goal. First, to estimate the number of defects remaining in the software, and second, to quantify the likelihood that these defects

result in system failures.

I'll talk about the procedure, the steps, using two examples. The first example is defect density. Defect density, defined as a ratio of unique defects found by inspections to the size of the product. The defects are classified into different criticality levels. And the inspections are requirement inspections, design inspections, and code inspection.

CHAIRMAN APOSTOLAKIS: How do you measure the size of the product?

DR. LI: The size can be either the source code size or the document size. The source code size can be the line of code, or it can be the function point. And the document size can be the number of pages, or it can be the number of paragraphs, or number of lines.

The effect of that, the requirement inspection, design inspection, and code inspection allow us to predict software quality at an early stage. Defect density has been widely accepted in the industry and academia. For instance, IEEE Standard 982.2 includes this measure. And the defect density is the de facto standard to measure software quality. A significant amount of research has been done using

this measure as a quality indicator.

MEMBER WHITE: I have a question about the size of the product. How do you handle the number of relationships that data have to other data, or that some line of software would have to data. In other words, I guess that's a complexity, actually, issue.

DR. LI: No, it's size, not complexity. They're different.

MEMBER WHITE: All right. So -- but you do take that into account then?

DR. LI: Right, right.

MEMBER WHITE: Okay, thank you.

CHAIRMAN APOSTOLAKIS: Well, I still don't understand. You say it's a de facto standard measure of quality. What is? You're doing a review of requirements and the code and all that, you identify the defects, and then you take that number, you divide by the size of the product?

DR. LI: Right, these are --

CHAIRMAN APOSTOLAKIS: What does that tell me now?

DR. LI: Well, that tells, you know, that -- it's the density. It tells how many defects potentially --

CHAIRMAN APOSTOLAKIS: I have found.

That's all it tells me.

DR. LI: Right, that have found. Right.

CHAIRMAN APOSTOLAKIS: So why is that a measure of quality?

DR. LI: Well, because the more you have the lower quality of your product. This measure historically --

CHAIRMAN APOSTOLAKIS: But again, wait a minute now. Are you applying this to a product that somebody tells you is ready to be used, or to a product that is in the process of being produced?

DR. LI: Sorry, I didn't get it completely.

CHAIRMAN APOSTOLAKIS: If it's part of the process, then you do find defects, because that's the whole idea. So are you doing it after the fact? In other words, now somebody has produced a product and says put it in your plant, and you go there, and you do a review, and you find a few errors.

DR. LI: Well --

CHAIRMAN APOSTOLAKIS: Is that what you mean?

DR. LI: Right. You can do both. In our institution, in our research right now we are doing, you know, the latter situation. We have a real

application from nuclear power plants. We have the entire document. And we are doing the inspection, from requirements to the code.

CHAIRMAN APOSTOLAKIS: Okay. So you find a particular number.

DR. LI: Right.

CHAIRMAN APOSTOLAKIS: And it certainly gives you an idea of how good it is, yes, I can't disagree with that. Sure.

DR. LI: Next we will quantify --

CHAIRMAN APOSTOLAKIS: You already have quantified.

DR. LI: -- the likelihood of these defects to the system failure.

CHAIRMAN APOSTOLAKIS: So far you have found the number of defects, and you divided by the size, and that's a number.

DR. LI: Right, that's a number.

CHAIRMAN APOSTOLAKIS: That's fine.

DR. LI: This is a standard. In other words, this is a measure found in the industry.

CHAIRMAN APOSTOLAKIS: Okay, let's go on and see now what you do with that number.

MEMBER GUARRO: One question.

DR. LI: Yes.

MEMBER GUARRO: What is your definitional defect in this context?

DR. LI: Well, we adopted IEEE definition, which is a deviation from the requirements. So all the terminology is here.

CHAIRMAN APOSTOLAKIS: So everything is compared to the requirements. If the requirements themselves are not self-consistent, what would that be?

DR. LI: Sorry?

CHAIRMAN APOSTOLAKIS: The requirements themselves are not a self-consistent set. Would that be a defect?

DR. LI: Right.

CHAIRMAN APOSTOLAKIS: Or you would never find it?

DR. LI: Well, we have specific measures to this --

CHAIRMAN APOSTOLAKIS: But that's not a deviation from the requirements. That's faulty requirements.

DR. LI: If there are any inconsistencies in the requirements, we have a specific measure to do that.

CHAIRMAN APOSTOLAKIS: But not this one.

DR. LI: Not this one.

CHAIRMAN APOSTOLAKIS: Okay. Okay.

DR. LI: We have certain measures.

MEMBER GUARRO: Okay, but also in -- just to pursue for a moment the issue here. Do you differentiate requirements in levels of criticality?

DR. LI: Yes.

MEMBER GUARRO: So you will classify defects also according to --

DR. LI: To the criticality level.

MEMBER GUARRO: -- the criticality level?

DR. LI: Yes.

MEMBER GUARRO: Okay.

MEMBER KRESS: And then what would you do with that classification? Would you put a weighting factor on the quantifier?

DR. LI: We have a specific technique so we can propagate this different criticality defect to the --

MEMBER KRESS: To the --

DR. LI: To the probability of failure. Because we can't review them differently. I will talk about shortly, you know, that special technique.

So given the value of defect density, then we can calculate the number of defects in the software

using this simple --

CHAIRMAN APOSTOLAKIS: Wait a minute, now. That's how you started. What do you mean you can calculate? You found them.

DR. LI: Right, right. We found the number of defects.

CHAIRMAN APOSTOLAKIS: Yes.

DR. LI: This assumes that if you have a defect density number provided by someone else, how you get to the number of the defects.

CHAIRMAN APOSTOLAKIS: This is a big step here. So you're saying `I found the DD in a particular program, and now somebody gives me another program.'

DR. LI: No, no, no. That's --

CHAIRMAN APOSTOLAKIS: I don't understand the situation.

DR. LI: There's two different situations here.

CHAIRMAN APOSTOLAKIS: Yes.

DR. LI: This relationship I just put here to highlight the relationship between the number of defects and defect density.

CHAIRMAN APOSTOLAKIS: Oh. So that's the definition of DD.

DR. LI: Right, right.

CHAIRMAN APOSTOLAKIS: All right.

DR. LI: So they found the standard, but the inspection may not find all the defects. The number of such latent defects can be estimated statistically using the capture/recapture techniques. Capture/recapture techniques were first applied in a study of the fish and wildlife populations. The simplest capture/recapture technique is a so-called two sample model. The first sample provided to individuals captured a mark that returned to the population, and the second sample provided the individuals recaptured. Using the number of individuals captured in both samples, and if the numbers captured is adjusted by one sample, one can estimate the number of not captured individuals, and then the entire population of the wildlife.

Recently, this technique has been applied in the software engineering field to estimate the number of defects not found by the inspection. In these applications, the number of defects is the analogy to the animal population size.

CHAIRMAN APOSTOLAKIS: Wait a minute. You are saying that you can estimate the population size from a small sample?

DR. LI: Right.

CHAIRMAN APOSTOLAKIS: Wow.

DR. LI: This technique has been --

CHAIRMAN APOSTOLAKIS: Don't you have to make some additional assumptions? I mean.

DR. LI: Right.

CHAIRMAN APOSTOLAKIS: So let's say I want to know how many coyotes there are in a particular place. What do I do? Capture a few and then extrapolate, or what?

DR. LI: Well, this is an entire discipline. And this technique has been validated for over 30 years in biology.

CHAIRMAN APOSTOLAKIS: Yes -- no. This is not an argument you can use here. You have to tell us why. You're asking me to believe somebody else. I have difficulty doing that. I don't understand how you can find five defects, and then you are able to tell me how many more there are. There's something missing there.

MR. CARTE: There's a couple of ways that this technique can be applied. One way, if you look at the animal population, you would choose a capture area that is representative of the total area. So in a software system, you would choose a set of modules

that are representative of all the modules in the system in terms of size, complexity, in terms of the different programmers or programming groups. So if you had a representative sample of modules that you applied this technique to, then you could estimate for the whole population.

CHAIRMAN APOSTOLAKIS: So there are additional assumptions, then. As you say, you go to an area that is more or less representative, and then you assume the density of animals is the same as in the bigger area.

MR. CARTE: Yes, that would be --

CHAIRMAN APOSTOLAKIS: Then I can understand how you can find that, but the question is whether these assumptions are valid.

MR. CARTE: Yes. That's one way that the measure can be applied. The other way that this measure could be applied, and that's why I mentioned licensee earlier, is if a licensee were to apply such a measure, they already have systems in place in terms of their QA procedures that completely review the entire system. They have multiple reviews in place. So if you used a capture/recapture model with removal, in other words once the defect is identified it's removed, and the multiple reviews, you can use these

techniques to estimate the number of defects remaining, and the type of defects, because you can categorize the type of defects found. So it can be applied in a complete system review by a licensee. I do not think that the NRC would be interested in having multiple reviewers do a complete review of the entire system of documentation. For that particular application, it is less likely to be done by the NRC, but reviewing a sample is more likely to be done. So in that sense it can be used.

CHAIRMAN APOSTOLAKIS: And that number of remaining defects can never become zero, can it? Because of the way you have structured the method? Which means now you have to tell NRR that if that number falls below a certain number it's acceptable.

MEMBER KRESS: It seems to me like this assumes you know the curve for the capture/recapture value versus the number of defects.

MR. CARTE: Well, the capture/recapture model, there's three methods of using defect density. There are in general three methods of using defect density to characterize remaining populations. One is capture/recapture, the other would be a neural network approach, and another would be the family of curve-fitting methods that you describe. But basically if

you have sufficient data, the equations behind capture/recapture are supposed to characterize the likelihood of capture of different types of defects because you have multiple reviewers and multiple capture rates. And so you can get estimates.

MEMBER KRESS: I can buy this. You do it several times and you get the start of a curve and extrapolate this curve.

MR. CARTE: Right. You have to have --

MEMBER KRESS: I don't see where a neural network comes into play.

MR. CARTE: Right. The idea with a neural network is that maybe these systems are non-linear, and neural networks do better at matching those.

MEMBER KRESS: See, it's just a way to correlate the data if it's non-linear.

MR. CARTE: Right.

CHAIRMAN APOSTOLAKIS: But are you going to tell us what to do with that number?

DR. LI: Yes. Next. Given the number of defects remaining in software, we utilize the so-called fault propagation technique to study the likelihood of these defects caught to the -- sorry, that the failure probability caught by this number of defects. And as the software engineering study has

shown, a defect will manifest itself as a failure if and only if the three following conditions are satisfied. First, this defect needs to be executed. Second, this defect needs to create a space anomaly. And the third, this state normally needs to propagate to the output of the software.

These three conditions are summarized in the PEI models proposed by Jeff Voas. And this is published in the 1990s in IEEE Transactions on Software Engineering. In these models, E represented the probability that a particular section of program is executed. I represented the probability that the execution of the execution of the problematic location affects the data state. And the P, the probability that an infection of the data state affects system output. Given the availability of P, I, and E, the software quality indicator, or the probability of failure per demand can be given using this equation.

Next, we utilize finite state machine techniques to quantify this model. Finite state machine models system behavior. This example models PIN entering function for a sample security gate system, which requires the entrant to enter the PIN. This model starts from the entry state, and at the end of the way the exit state. A rectangle represents a

state. An arc represents a tradition. A link from entry to exit constitutes a task. The probability of each transition are embedded in the finite state machines. So the probability of each task can be calculated as a product of the probability of each transition within that task.

CHAIRMAN APOSTOLAKIS: Can you give us an example?

DR. LI: Yes. For instance here, you have the -- from the start, you need to enter the PIN. The PIN, you have two conditions. One is a good PIN, and the other one is a bad PIN. So the probability of the good PIN can be 0.8, and the probability of the bad PIN can be 0.2.

CHAIRMAN APOSTOLAKIS: Why?

DR. LI: This data is from the user profile, from the log file. We obtain this data from the field data, from this profile from the field data.

CHAIRMAN APOSTOLAKIS: And?

DR. LI: Then we map the defects to this model. And this dashed line shows the defects located here. Then we know the task that travels this transition will lead to a failure. So the integral of the probability of the task that travels this transition will provide us the estimation of the

probability of failure caused by this defect. We repeat this procedure for all defects. Then we have the overall probability of failure per demand.

CHAIRMAN APOSTOLAKIS: And for all defects you will have this information of 0.8 versus 0.2?

DR. LI: Right, right.

CHAIRMAN APOSTOLAKIS: I can't see how. I mean, this was a very concrete example. You know, you can go there and type in their PIN, and they make a mistake. And you know that, and you can find it. But what if you have something esoteric, somewhere there buried. I mean I don't know how --

DR. LI: Well, let's talk about the actual --

CHAIRMAN APOSTOLAKIS: You know the probability of each path. Wow. That's a pretty strong statement, isn't it? Because that assumes that all these probabilities are external, aren't they?

DR. LI: Well, currently --

CHAIRMAN APOSTOLAKIS: That would be which probability that you showed us earlier, P?

DR. LI: That's P.

CHAIRMAN APOSTOLAKIS: Okay.

DR. LI: Oh, sorry.

CHAIRMAN APOSTOLAKIS: That's P?

DR. LI: No, that's E. Execution.

CHAIRMAN APOSTOLAKIS: Okay.

DR. LI: And with the way we build this finite state machine, we can guarantee that E and I are equal to 1.

CHAIRMAN APOSTOLAKIS: So if I have now a reactor protection system, it's monitoring a fairly large number of parameters, you will be -- wouldn't E be the probability of any possible combination of values of these?

DR. LI: That's correct.

CHAIRMAN APOSTOLAKIS: And you will know what the probability of these combinations is?

DR. LI: Yes. Currently --

CHAIRMAN APOSTOLAKIS: How on earth would you know?

DR. LI: Currently we have the data from the actual nuclear power plant.

CHAIRMAN APOSTOLAKIS: How would you know?

DR. LI: They maintain a comprehensive log data, data file.

CHAIRMAN APOSTOLAKIS: No, but I'm talking about accidents here. I'm not talking about normal operations.

DR. LI: Yes, that's what I'm talking

about.

CHAIRMAN APOSTOLAKIS: How many accidents have we had? So that we'll be able to say the probability of this combination of variable values is that. I don't see how we can know that. I mean, you can have weird situations where you have to SCRAM. And you're saying, no, I will know the probability that I will have this weird combination. Maybe you do, but I have to be convinced a little more.

MEMBER GUARRO: Well, this brings back a point that was, I think in a previous chart there was as an indicator of quality was mean time to failure. Mean time to failure is something you can measure in a system that you operate normally. You can observe and recover from failures. But when you're talking about severe accidents, mean time to failure is something that doesn't mean much as an indicator of performance, because you don't see mean time to failure as measurable, right? So this is an important point to keep in mind when translating statistics taken from a routine type of application, extrapolating to a rare accident scenario type of application.

CHAIRMAN APOSTOLAKIS: Yes. And how would this apply to the examples, who did it this morning,

I think it was Mike Waterman. The Turkey Point and Davis-Besse, real incidents. Would you take -- not right now -- would you take your model and go to that piece of operating experience and tell us how you would have predicted that? How would you have assigned a probability to this problem with the sequencers? I think it's awfully hard. I mean, it's one thing to talk about people typing in personal identification numbers, and quite another dealing with a nuclear reactor.

MR. ARNDT: There's two issues here, both of which are important, but have different aspects. One is, as rightly pointed out, your operational profile of how these finite state machines work, and where they go, and things like that, it's difficult to get a complete characterization because, as you get to lower and lower probability events it's harder and harder to predict those. The other issue is predicting by some kind of analysis methodology this one or anything else, interactions that exist, failures or whatever, that you just haven't thought about. By characterizing in a more formalized way the analysis of particular kinds of things. In this case, if you write the detailed state space evaluation of the system, you then have something to hang onto, and

you can look at them in a formalized way. So without actually doing the analysis, I would say likely we would have caught things like the Turkey Point analysis, because we just didn't look at it, because we didn't have a formalized, organized way to look at it. Both of those are very valid points, but they're different issues.

CHAIRMAN APOSTOLAKIS: Yes. But this is not being advertised as being a methodology that helps you look at the structure of the software. It's advertised as a methodology that produces a probability. And it would be critiqued as such. I mean, I fully appreciate that, you know, I mean the standard -- if you do a full tree analysis, you really understand your system independently of how good your numbers at the end are.

MR. ARNDT: Yes. And what Bill and I tried to point out in the earlier presentation is that the programs under the software quality assurance program have multiple roles. The primary role is to better understand the system, and secondarily have more quantitative assistant approaches to do that.

CHAIRMAN APOSTOLAKIS: And a number of methodologies out there deal with the internal workings of a system.

MR. ARNDT: Right.

CHAIRMAN APOSTOLAKIS: I have no problem presenting them as such. I do appreciate that you're learning a lot about the system by attempting to do this, and other things. Maybe that should be a project. But when you start saying that I will calculate the probability by taking this integral, and I will need E, P, and whatever else it is, I just don't know that you can do it, Mr. Li. I really want to believe you, but I cannot. So try to convince me. I'm really on your side. I just can't accept this. I think it's too optimistic. I have to be frank with you.

DR. LI: I think the best way to convince is to wait for us to finish our real application.

CHAIRMAN APOSTOLAKIS: Then it's no fun if I wait.

MR. KEMPER: That's what I was going to suggest. This is Bill Kemper again. Perhaps if you'd like --

CHAIRMAN APOSTOLAKIS: Okay.

MR. KEMPER: -- we can certainly dove into this when we get close to the endpoint and provide whatever exposure you need, George, to the process.

CHAIRMAN APOSTOLAKIS: No, I'm not saying

that you should stop. I'm just giving you my problems as you go.

DR. LI: Yes, I understand that.

CHAIRMAN APOSTOLAKIS: But you also have to understand that giving an example with somebody typing in a PIN is not a very convincing argument. You're talking to Advisory Committee Reactor Safeguards. I mean, we don't care what people do when they type their PINs.

DR. LI: There's another entire discipline to study how to obtain --

CHAIRMAN APOSTOLAKIS: You have to immediately think in terms of safety.

MR. ARNDT: Right.

DR. LI: Correct.

MR. ARNDT: And that was one of the critiques that we got on the preliminary evaluation was that it needs to be a system designed to be implemented in a nuclear environment, which is why we're using a different nuclear system --

CHAIRMAN APOSTOLAKIS: Okay.

MR. ARNDT: -- for the secondary evaluation. Go ahead.

DR. LI: My next example is statement test coverage. Statement test coverage, defined as a

portion of software statements executed against a set of test cases. This measure is also --

CHAIRMAN APOSTOLAKIS: So excuse me, now you are trying to figure out what P is, is that correct?

MR. ARNDT: This is a different measure.

DR. LI: That's another measure. Sorry.

CHAIRMAN APOSTOLAKIS: Oh.

DR. LI: It's on Page 14.

CHAIRMAN APOSTOLAKIS: I know. But did you tell us how we would get the other probabilities? Like P and I?

DR. LI: Oh. Well, just as I discussed, P and I are equal to 1. You know, the way to develop this finite state machine model can guarantee that P and I are equal to 1. If P and I are not equal to 1, which means there are conditions keep the defect from being infected and propagated. So in the finite state machine model, you should be able to decompose and to identify, the describe these conditions. Just like additional branches. So the advantage of this finite state machine model technique is that you reduce the PIE model to the E model.

My next example is test coverage, the statement test coverage. The statement test coverage

is defined as the software statements executed against a set of test cases. This measure has also been widely accepted in the industry and academia. The IEEE standard also includes this measure. And this measure is commonly used in the software industry to control testing process. In particular, Malaiya studied the relationship between the defect density and the number -- sorry, test coverage and the number of defects. And this slide summarizes such quantitative relationship. This is empirical relationship. C_1 is a statement test coverage. And C_0 is the intermediate result which represented the portion of the defects found by the testing. And A_0, offer 0 to offer 1 are coefficients. And the N_0 is the number of defects found in testing. So N is the number of defects remaining.

CHAIRMAN APOSTOLAKIS: C_0 is what, defect calculation?

DR. LI: Defect coverage, which is the portion of defects found by testing. N_0 is the number of defects found by testing.

CHAIRMAN APOSTOLAKIS: No. Coverage means the portion of statements executed.

DR. LI: That's C_1. It's called test coverage, statement coverage. C_0 is defect coverage.

This is the intermediate result.

CHAIRMAN APOSTOLAKIS: What is the justification of this logarithmic exponential equation?

DR. LI: Well, this work --

CHAIRMAN APOSTOLAKIS: Is it a vehicle or what?

DR. LI: This is an empirical -- well, I will say coefficient relationship. This one published in the International Symposium on Software Engineering Conference. And we validated this relationship using two applications which are summarized in NUREG-6848.

CHAIRMAN APOSTOLAKIS: Validated.

DR. LI: And again, we utilize finite state machine techniques to quantify --

CHAIRMAN APOSTOLAKIS: You know, our handouts don't have the equation. Why? We have blanks.

MR. CARTE: That's an editorial problem on my part. They're there, they're just printed in the color white.

(Laughter)

CHAIRMAN APOSTOLAKIS: White characters on white background. There was a play that won the Pulitzer Prize. It was about a painting that was

white stripes and white background.

I also, I'm uncomfortable when you say it's used widely by the industry. I mean, our staff went and talked to the industry in the `90s, and the message was don't go near those methods. Not just these, any methods. So now you're saying they're used widely? Maybe that's a slight exaggeration on your part? I mean, does Boeing use things like that? Does Airbus use them? I doubt it. And you know, there was a paper in a conference, yes sure, as you know there are many papers in many conferences.

DR. LI: You mean the measure itself --

CHAIRMAN APOSTOLAKIS: Yes.

DR. LI: -- it's relationship.

CHAIRMAN APOSTOLAKIS: Yes. I mean, do you know of any serious industry that's really using it and makes decisions using that?

MR. ARNDT: George, part of the issue is a lot of the metrics are used, but exactly what they're used for is really the more appropriate question. Using metrics to improve the development process was the original intent, to, all right, are we getting enough coverage, are we finding enough faults, should we ship a product based on X. Part of the -- the whole purpose of this research is can you use

metrics that are used in the design process in the assessment process.

CHAIRMAN APOSTOLAKIS: I think you're going to have a major problem with this estimation of the number of defects remaining in that you will have to eventually tell us what's acceptable. And I don't know how NRR can approve something knowing that there is a number of defects remaining. On the other hand, you might say we are licensing reactors, so we know there's a probability of a major accident. I don't know, guys. The thing obviously leaves me very uncomfortable. But again, I'm willing to be convinced.

MR. ARNDT: One of the other issues is we don't have to use this as a strict quantifiable, go/no-go decision. If we, at the end of the research, at the end of the current project we're looking at, which is trying to validate the methodologies for a larger system, the result may be quantitative go/no-go decisions are not possible. However, the use of the various families of metrics, ones that look at complexity versus ones that look at other things will give us an indication of where in the system there may be bigger problems. The system may be exhibiting too much complexity, it's driving the number up relative

to the other metrics, therefore you should spend more time looking at complexity issues. So the point here is we're trying to figure out how much of this can we use in a regulatory environment. I mean, if the project succeeds wildly beyond our dreams, then we could maybe get to the point of quantification for a go/no-go, but that doesn't necessarily mean that's how we're going to use it.

CHAIRMAN APOSTOLAKIS: The other philosophical objection I have is that it focuses so much on the number of defects. If you come from the, you know, safety perspective, the number if probably relevant, but really it's the quality. It's the kinds of defects that I have. That scares me much more than just the number. And this seems to be focusing exclusively on numbers.

And you know, coming back to Dr. Guarro's question, how do you define the defect? You said the violation of the requirements. Well, that's pretty general. But --

MEMBER GUARRO: That could be something when the screen comes the color yellow instead of blue.

CHAIRMAN APOSTOLAKIS: Yes. And I have a thousand of those. I don't care.

MEMBER GUARRO: It should be blue, and then you define it.

CHAIRMAN APOSTOLAKIS: Like you know, the type of the equation is in white. That's a defect. But I don't care. We can fix it. It's not a safety-related defect. I'm interested in the safety-related defect. And I don't see how this can find it. What if you say, okay, you have coverage, right? And you find -- in the previous one, defect density, right? Tell me, what is a typical number of defects one finds? Eleven? I don't know. A hundred? Seventy-two?

MR. KEMPER: Slide 18, I think, is where.

CHAIRMAN APOSTOLAKIS: 18?

MR. KEMPER: In the next few slides we'll give you some numbers, but the point I wanted to try to make though is -- yes, Slide 18, we've got some numbers ahead of you. We're going to talk to you about. But the point I was trying to make here, these metrics -- we've already said it before. It cannot replace the human being, the human element.

CHAIRMAN APOSTOLAKIS: Yes.

MR. KEMPER: In other words, the idea is these hopefully will be a pointer for experienced, seasoned reviewers to help them assess where they

should focus their detailed review.

CHAIRMAN APOSTOLAKIS: But again, I don't know. It seems we are going out of our way to find something useful here. Because you say, I mean on Slide 18 it says 210 were highly ranked measures, and so on. What if one of these 210 is failure? Failure. You have core meltdown, and the whole thing. I mean, I wouldn't put it as 1 out of 210. I would say this is really the real deal, I have to look at it, and understand it, and eliminate it. And these methods don't do that. They look at numbers.

DR. LI: Well, the fact is that we do look at the criticality. We do look at the effect of different defects.

CHAIRMAN APOSTOLAKIS: And then what do you do with them, though? You don't seem to do much about them.

DR. LI: Just like I mentioned in this diagram, in order to map that defect to this model, you have to understand semantically what does that defect mean. What the defect --

CHAIRMAN APOSTOLAKIS: I know that. But then you go on and calculate densities, you calculate C_1, C_0, and so on. The severity enters in a very crude way in your classification of criticality.

DR. LI: Right, but defect density does count the criticality. Test coverage is a different measure. That's why we have different measures.

CHAIRMAN APOSTOLAKIS: Let me ask you something else. Are these gentlemen, or ladies, Pfleeger, Malaiya, are they working on high consequence industries? Or are they working on PCs? I mean, do they worry about severe consequences in their software evaluation?

DR. LI: I will say they are software engineering people.

CHAIRMAN APOSTOLAKIS: So they don't get --

(Laughter)

DR. LI: They work at Microsoft.

CHAIRMAN APOSTOLAKIS: Well, I mean, yes. If your biggest worry is that Microsoft Word works most of the time, it seems to me you have a certain number of concerns. And if you don't want to have radioactivity release, you have another number of concerns. Very different approaches. Very different mindsets.

MR. ARNDT: There's been work in all parts of the software engineering community. And that's actually one of the biggest challenges in some of

these areas, particularly the ones where the quantification, or the analysis is based on empirical curves, is to determine whether or not that particular empirical curve is sufficiently based in safety-related applications, or is it just a compilation of everything. So that those particular issues are things that we're trying to attack at the various points.

CHAIRMAN APOSTOLAKIS: Let me understand something else now here. This session is supposed to go until 2:30. Is your presentation going to be until 2:30, or there's more?

MR. CARTE: I have two slides when he's done.

CHAIRMAN APOSTOLAKIS: Okay, okay. So we're doing fine. So can you go to 18?

DR. LI: 18?

CHAIRMAN APOSTOLAKIS: Well, or no here, 17.

DR. LI: 17. Okay, this slide summarizes the current status. And we apply 12 measures to a real nuclear application. It's an I&C application. And the measurement in progress and their completion date, summarized in this table. And the further analysis required. By July 15 we need to build up the

operational profile. By August 15, we need to build the finite state machine. And by August 15, we need to perform a reliability testing. And this -- the final analysis needs to be done by September 30.

CHAIRMAN APOSTOLAKIS: Is that when the contract ends?

MR. ARNDT: No, the contract goes till November to get the report finished.

CHAIRMAN APOSTOLAKIS: Now, last time that you guys were here from Maryland, you told us about how you surveyed experts, and they told you, you know, how, what is the conditional probability that this measure gives you a good idea as to how good the program is. Am I saying it correctly?

DR. LI: Well, basically the expert opinion elicitation study provide an indicator about which measure is better in terms of predicting software quality. So that's one --

CHAIRMAN APOSTOLAKIS: That was in addition to this.

DR. LI: Sorry?

CHAIRMAN APOSTOLAKIS: It was in addition.

DR. LI: Right.

MR. ARNDT: It was an input to this program.

CHAIRMAN APOSTOLAKIS: If it's input then I want to understand how. If -- again, you know, a great way of answering my questions is to put yourself in the shoes of the regulatory staff this morning. They receive this application from Oconee. How would you apply your method to help them make a decision? If you give them a generic statement, like the defect density according to the experts is a good indicator 36 percent of the time, I just don't know what they can do with that. Because they are dealing with a specific system. If you can give them more specific information, then more power to you, great. This is really the test, not that somebody presented a paper in 1994. So they have this issue in their hands. How could something like this be helpful to the decision-maker?

MR. CARTE: There are a couple of ways that this could be helpful to the decision-maker. One, if the licensee implements a measurement program, then the NRC could review the measurement program and use that to increase their level of assurance that the system provided is okay. One of the things that Steve mentioned earlier is that this research stems from the design engineering research. So basically, when you look at the IEEE standards regarding measurement, they

all talk about establishing a measurement program. And in some of the literature it talks about you have to wait a few years before you really see the results of the measurement program. And they are based on a stable process. So given a stable design process, you are able to characterize, or statistically make characterizations about the product. So one application is that if a licensee implements a measurement program, and implements it correctly, that can give us reassurance, and allow us the possibility to look at a smaller sample of threat audits.

I mean, if we're doing a sample of threat audits, those should be statistically characterizable of the system in general. Can we look at a smaller number of audits. Can we rely on the measurements that they use. And that's part of -- to understand how good these measurements are. If they give us -- we've both done measurements, but -- and then we look at those measurements, we need to have some assurance, or some confidence that measurement programs and the types of measurements are actually useful in predicting or indicating reliability or quality. It's more difficult to implement a measure on a piece of software that arrives. Defect density is a measure that could be done, in a sense, but what that would

require is at least two reviewers to review a product. And that sample product would be representative. And then from that you could calculate how many latent defects there are. You could also characterize the type of defects there are. And you basically get latent defects from defects found by one reviewer and not the others. So that indicates that these defects are not as easily encounterable.

And when you talk about quality, there are many dimensions of software quality, and maintainability is one of them. How cohesive are the specifications, how modular are the specifications. The same rules that you apply to source code review can be applied to document review, in terms of cohesiveness, clarity, modularity. So not all the defects identified are -- will impact the proper functioning of the system.

MR. ARNDT: The point is we're trying to understand whether or not methods like this are usable. And if you go out and try and use them in a test case --

CHAIRMAN APOSTOLAKIS: I am trying too, Steve.

MR. ARNDT: Okay.

CHAIRMAN APOSTOLAKIS: I really am trying

myself. But I seem to be a little more skeptical than you are.

MR. ARNDT: Fair enough.

CHAIRMAN APOSTOLAKIS: Which is fine.

MR. KEMPER: If I could offer one thought too, just to kind of tag onto what Steven just said. You know, we're -- this project is a three-phase project as you're aware, and that we're really trying to assess the viability of these metrics on a complex system using nuclear power plants. Actual deployment of this technology now into inspection criteria is a yet-to-be-determined project. So we'll build onto the results of this to actually figure out how to actually implement this into the regulatory process.

CHAIRMAN APOSTOLAKIS: If --

MR. KEMPER: If it's useful, yes, exactly.

CHAIRMAN APOSTOLAKIS: Well, fine. You know, I have no problem with that.

DR. LI: This slide summarizes our preliminary results so far that we obtained. The number of defects predicted from the completed measures.

CHAIRMAN APOSTOLAKIS: Which program are you applying this to now?

DR. LI: It's a real nuclear software. I

cannot disclose the name of the software based on the agreement with the vendor.

CHAIRMAN APOSTOLAKIS: You found 210 highly ranked?

DR. LI: I just tried to highlight that number, 210, from cyclomatic complexity. It's not the number of defects remaining. It's the number of defects before the testing. So ongoing research is trying to explore how many defects are remaining.

CHAIRMAN APOSTOLAKIS: Okay.

DR. LI: And another point is that bugs per line of code. This measure is obsolete. So the value from that measure is not representative.

CHAIRMAN APOSTOLAKIS: Which one is obsolete?

DR. LI: The bugs per line of code. Bugs per LOC here.

CHAIRMAN APOSTOLAKIS: Oh.

DR. LI: And I also want to highlight that although the measure cause effect graphing ranked by the experts in low category, but the way we measure, it significantly promotes the ranking of this measure. So that's why we have a very low number of defects predicted from this.

MEMBER WHITE: Excuse me. Can you tell me

why bugs per line of code is irrelevant, or unimportant?

DR. LI: Well, this measure was invented in the 1970s, and based on the data from assembly language. So the line of code for the high-level language like C, and from the low language like Assembler, are significantly different. So that's why this measure and this empirical relation between the number of bugs and the line of code.

MEMBER WHITE: I understand that argument, but the number 590 is still pretty large.

DR. LI: Right.

MEMBER WHITE: And so that would cause me, you know, to -- it would cause me some anxiety. So why would we still -- why would we consider that irrelevant? I understand about lines of code, but 590 is a big number, right?

DR. LI: Right. Well, that's why the experts rank this measure very low. So which indicates that everybody should not take this measure.

MEMBER WHITE: You'll help me, won't you.

MR. ARNDT: What you've got to realize is one of the purposes of doing a validation study is to try and determine which measures may be useful, and are predictive of what the reality is. So what the

idea is is to look at some of the low ranked measures, ones that we have less going in confidence that will be useful, to, one, validate that that's true, and, two, also decide that, yes, we don't hold a lot of confidence in that particular measure even though it's out there in the community. And that if a licensee at some point in the future says, well, you guys are interested in metrics, I'll throw this into my application, we can say, well, that's nice, but based on our research it's pretty useless. So the point is that we want to look at a variety of measures to understand not only how easy are they to use, what information do they give you from an understanding of the system, but also whether or not we would add any value to them in a licensing review. So the idea is to look at a number of different issues.

What Ming was pointing out is in some cases it depends on how the metric is defined. In this case, it's not well defined anymore based on -- because we don't program in Assembler very much anymore. Other cases like cause effect graphing depends on how well the procedure for developing that metric is defined. As Ming mentioned earlier, as part of this research we better defined that procedure, so we now believe it is probably a higher ranked measure

because the consistency in developing that measure is higher than it was when we first started looking at it, and ranked it as a low ranked measure.

MEMBER WHITE: Okay. Under medium ranked measurements you have, is that capability maturity model?

DR. LI: Right.

MEMBER WHITE: And the 4.58 is between 4 and 5. But that's a medium ranked measure. And the cyclomatic complexity is a high ranked measure? What does the number 210 mean?

DR. LI: Well, just as I mentioned, this is not the number of defects remaining. This is the number of defects before testing. So after the testing, the development process will fix most of the defects here. So this is just a preliminary result. And we are working on that, try to theoretically figure out how many defects are remaining.

MEMBER WHITE: Okay.

MEMBER GUARRO: I'm having some trouble in relating the concept of number of defects to these measures, actually. For example, in cyclomatic complexity, what 210.37 means. Some metric? Because the label says number of defects, and I'm not sure --

CHAIRMAN APOSTOLAKIS: It says predicted.

MEMBER GUARRO: Well, if I interpret that correctly, in bugs per line of code I'm going to have 590 bugs per line of code?

CHAIRMAN APOSTOLAKIS: That's crazy.

MEMBER GUARRO: That doesn't seem to be the meaning of what you have there.

MR. ARNDT: Well, let me do the simple answer, and Ming can elaborate the more complicated. What we're trying to do so we can make a comparison on relative value is we're getting the actual number out of whatever the particular metric is, and then we're using published literature, or correlations, or whatever for each different measure to try and normalize each of the measures to a particular value, like number of defects predicted, or some other normalized value. That's what those numbers are.

MEMBER WHITE: Since we have a little time, and since I'm an old country boy, maybe you could help me a little bit more. If I'm from the NRR, and you tell me that this safety-related application, digital system does have a normalized value of whatever it is, let's say it's 210. What does that tell me? How do I use that information? What do I do with it? Does that tell me it's good code, bad code, I ought to be worried about it, I ought to throw it

out? I'm sorry if I'm dense, I just don't understand yet.

DR. LI: This is not the final result. The final result we will provide the probability of failure per demand. By combining the operational profile and the number of --

CHAIRMAN APOSTOLAKIS: Any demand? What do you mean probability of failure per demand? This is conditional probability. Depends on the demand.

DR. LI: The system we're studying is an RPS system. So by "demand" we mean it's a per trip.

CHAIRMAN APOSTOLAKIS: Well, even so, I mean any combination of variables will give me the probability of failure? Okay, go ahead, then what? Then you will provide that probability which will be what? 0.02, something like that?

DR. LI: Well, we don't know the results yet.

CHAIRMAN APOSTOLAKIS: But let's say it's 0.02. The question from Mr. White is what do you do with that.

DR. LI: Well, from the software quality perspective, that value tells us if you run it one hundred times, you will experience two failures.

CHAIRMAN APOSTOLAKIS: Yes.

DR. LI: That's a statistical indicator.

MR. CARTE: We're talking about measures and results that they produce, but we have not established acceptance criteria. That's the point where you establish whether the result produced is acceptable or not.

CHAIRMAN APOSTOLAKIS: Well, the other question is of course whether the probability should be 0.02, or you should have some sort of an uncertainty range associated with that.

DR. LI: Well, that's in our next step.

CHAIRMAN APOSTOLAKIS: I think you're a brave man to claim that you will produce a probability of failure based on these measures. I am very, very skeptical. Anyway, let's keep going. 19 is your future?

MR. CARTE: Yes. So the future work in the large-scale validation will in part include the development of -- first we have to determine which methods are acceptable. And from that we can look at what is the acceptance criteria. And there's a couple of ways of developing acceptance criteria. And one is to apply these measures -- which is called benchmarking -- one is to apply the measure in parallel with the current evaluation process, and

based on what we currently deem as acceptable, what are the measures of that software. And that gives us a relative estimate of the acceptance criteria that we should look for. So in future when systems come in and their measures are significantly below that, the currently acceptable levels, that should cause some concern.

The other aspect of acceptance criteria relates to this calculation of failure per demand, probability of failure per demand. And one reason to pursue a method like that is that it gives you a theoretical way of determining an acceptance criteria. If you can characterize, at least statistically, what you estimate the failure probability to be, then you could apply PRAs and from that get an acceptance criteria. If that works, that is less work than benchmarking, because how long do you have to benchmark a measure before you have confidence in that measure? And so, yes it is a little bit cutting edge to pursue that, but that's part of the motivation for pursuing it.

CHAIRMAN APOSTOLAKIS: When you say future work, you mean after Maryland finishes in November? Or future in the next few months?

MR. CARTE: Well, both. The first step,

because the research isn't done, the research final report will include an evaluation about the acceptability of those measures. And that will be reviewed by us, and our NUREGs are circulated to NRR for review. And so we'll evaluate that. If those methods are deemed to be acceptable by the NRC, then we will need to look at training and the use of those metrics. If they're not acceptable, then that work is in essence done. If we the NRC, and that includes input from NRR, determine that this is still promising and we wish to look at additional measures, we can pursue that as subsequent research. And another area of subsequent research is technology-specific measures. For instance, right now there are three SERs for PLC-based systems, and yet we're not looking at PLC-specific measures. How does lines of code apply to a function block design, for instance?

Basically we feel that software engineering measures are sufficiently mature for assessing software quality in safety-related nuclear applications.

CHAIRMAN APOSTOLAKIS: I thought the comments from both you and Steve so far pointed to the conclusion that you're really not sure. But now you're definitive. I thought you were still

exploring, and now you're saying no, they're sufficiently matured.

MR. CARTE: They've matured for performing assessment, yes. Whether we have an absolute acceptance criteria, or how we use those numbers -- a quantitative assessment gives you more granularity in the performance of your review. Also, if you have detailed measurement rules it gives you a more defined process.

CHAIRMAN APOSTOLAKIS: I agree with all -- these are generic statements. In this particular approach it seems to me you have to really scrutinize, like in any approach, the fundamental assumptions. And the problem with software is that, as someone said this morning, there's usually specification errors, design requirement errors, and so on. And 99.9 if not 100 percent of the matters we have here really do not apply. We don't deal with those kinds of errors in standard risk assessments. So we really have to go back to the assumptions, every step of the way. You know, they say this, I can say something about the remaining faults. No. For me, that's a major claim. It requires major arguments. I don't see them. So I must say at this point I disagree with the first bullet. That doesn't mean you shouldn't agree with

it. I mean, in the future sometime we have to resolve this. I'm awfully skeptical about all this. I really don't think it gives you anything. There you are. But then again, I may be wrong. Right? We'll find my P and my E integrate. So, the last bullet says what now, Norbert?

MR. CARTE: Measures of software quality are related to proper system operation. And this large-scale validation project provides a promising methodology for estimating the impact of software quality on proper system operation.

CHAIRMAN APOSTOLAKIS: Okay. Is your presentation over? Any questions? Comments?

MEMBER WHITE: I have a question. This candidate system that you're evaluating in your project, what was the requirement for reliability? Was it like one failure in 10^{-6}, or 1 in 10^{-4}, 10^{-2}?

DR. LI: These were not mentioned explicitly in the requirements.

MEMBER WHITE: Okay. Well, the point is what you're using in your project is a highly -- is supposed to be a highly reliable system, right? But you can't characterize exactly what that is right now. But it's like -- it's better than 1 in 100? One in 1,000?

MR. ARNDT: Although, to my understanding, and I could be wrong, but to my understanding the actual line criteria was not specified when it was originally designed. If you go back to the standards that it does reference in its design work, you can infer based on some other standards 10^{-4}, 10^{-5} ballpark.

MEMBER WHITE: Thank you.

MEMBER GUARRO: Can you go back to Slide 15 so we can write in the formulas?

CHAIRMAN APOSTOLAKIS: What? Oh I think Eric is doing that. Sergio? He's going to do it. Okay, any more comments or questions? There is a question here.

MEMBER BONACA: I was missing the first half an hour. I had a meeting here. But I just, on reviewing this report here on preliminary validation as a NUREG. I was intrigued by, again, you had the Table 1 on Page 7 where you identify 40 or 30-odd measures. And you pick up two high ranking class, two medium, two low. You work with those. It draws out the conclusion, and then you seem to be able to apply those conclusions to the whole set.

MR. CARTE: We get some indication of the validity of the ranking. In other words, for those

measures that are low ranked, do they perform low ranked for those that are medium ranks.

MEMBER BONACA: That was a specific purpose.

MR. CARTE: Yes.

MEMBER BONACA: In fact you had some changes in rank that resulted from the evaluation.

MR. CARTE: Yes.

MEMBER BONACA: Okay. Okay.

CHAIRMAN APOSTOLAKIS: Okay. Anything else?

MR. WATERMAN: Mike Waterman, Research. Just from an NRR perspective, can't get that out of my blood, I guess. On Slide 18 where you showed the preliminary results, and you've got number of defects predicted. Have you considered building a system where you actually knew how many defects were in the system so that you could check out and see just how well these particular metrics, for example, were predicting defects when you already knew the answer? I don't see a benchmark -- I don't know if there were actually 4.58 defects remaining in the system, or if there are 200 by looking at this chart. All I see is the numbers, and you don't have anything to weigh those numbers against, you know, what is really in the

system. And I think that would be very helpful, you know. Because right now none of those numbers mean anything to me other than if I was using bugs per line of code, and I was an NRR reviewer, I'd get pretty excited pretty quick. And I'd know that I'd have to extend an audit by several weeks just to chew into that. So right now I'm, just from my experience as a reviewer, those numbers there sort of disturb me unless I know how many defects are there really remaining. Then I could say, oh yes, cyclomatic complexity, how ridiculous. And look, CMM does a pretty good job. You know, I don't know that by looking at that. So it would seem to me somewhere down Research's road there would be a benchmark model where you know all the answers. You apply these things to that benchmark model, and see how well it does in finding the right answer. I don't know if that's in the research or not. That's Norbert's research project.

DR. LI: Right, this is absolutely correct. We will do a reliability testing later. And based on that reliability testing, we will know how many defects are really remaining in the system.

MR. WATERMAN: But it seems to me you'd have to find every defect so that you could see how

well you come out on predicting number of defects, right?

DR. LI: Right.

MR. WATERMAN: And then I don't know how reliability relates to defects if you have a defect that doesn't affect reliability.

MR. ARNDT: Right. Well, that's the difference between a failure and a defect.

MR. WATERMAN: Yes.

MR. ARNDT: If you look at the slide before the one that's right up there, if you go up, 17. You look at the analysis and progress. Part of the effort is to do some testing to get -- for the system under consideration to get a failure on demand estimate to validate the predictions that the metrics will provide you.

MR. WATERMAN: Well, could we use --

MR. ARNDT: -- a rough evaluation of whether or not the predictions are reasonable, and which metrics are most closely tied to the test base prediction.

MR. WATERMAN: Well, Roman Shaffer from my section made a suggestion I thought was pretty ingenuous, was to take our fault injection tool that we've got, and apply it to your benchmark model, and

let the fault injection shake the daylights out of it, if you will, and see how many bugs fall out, and then use that as a benchmark against all of these things.

CHAIRMAN APOSTOLAKIS: Could be.

DR. LI: Yes, that's possible.

MR. ARNDT: There's a number of different methodologies for trying to get a reasonable prediction based on a different methodology to support which metrics are the most accurate.

MR. WATERMAN: And I guess finally, as a reviewer of a system, having a large number of metrics would probably really assist me because they would point me in directions that I needed to go when I actually reviewed the product manually, instead of just relying on just these numbers. I would hope that the reviewers who were remaining in NRR would use those numbers to tunnel down in to very certain aspects of a particular product and see why that particular aspect isn't coming out so great. So you know, so I look at this research as kind of helpful in that way. That's all.

CHAIRMAN APOSTOLAKIS: Thank you very much, gentlemen.

MR. ARNDT: Thank you.

CHAIRMAN APOSTOLAKIS: And we'll recess

until 2:45.

(Whereupon, the foregoing matter went off the record at 2:18 p.m. and went back on the record at 2:45 p.m.).

CHAIRMAN APOSTOLAKIS: Back in session. Mr. Arndt?

MR. ARNDT: Yes, sir.

CHAIRMAN APOSTOLAKIS: The floor is yours.

MR. ARNDT: Okay. We're going to talk now a little bit about the project that is identified under Section 3.2.2 in the Research Program Plan. This is the digital system dependability. Myself, who you all know, and Mr. Shaffer will give this presentation. I'll just do the brief introduction, and then Roman will do the meat of the presentation. I will of course be available for questions.

As we talked about this early afternoon, this is part of the software quality assurance program. And this part of the overall program is designed to look at different testing aspects to understand digital system dependability in a more detailed fashion. Next slide, please.

As we talked about this morning, the current state-of-the-art for these various digital systems includes a very promising methodology referred

to as fault injection testing that permits the system to be reviewed at a fairly deep level. Once you obtain the information, or a better understanding of how the system works, that could then support modeling methodologies in a number of different ways. And it doesn't really matter which modeling methodology you use to embed the information you learn about the system. The idea here is to characterize the behavior of the system using this particular methodology. In this case, although fault injection has been historically looked at in the software area, there's also been work in the hardware area, in the total digital system area for integrated hardware/software interactions. People have done it in the simulation-based arena as well. So there's a number of different ways you can do this. We're going to look at it in a particular way to try and develop a better understanding of the system. So the idea here is to develop an understanding of the various aspects of how the system can fail, and information we can gain out of these kinds of techniques.

Roman is going to give you some more details of what the specific goals are for this project. This project basically is an out-cropping of information we gained under a cooperative agreement

with the University of Virginia and other research programs. We now want to take that information that we gained and use it to develop specific applications.

MR. SHAFFER: Thanks. Good afternoon. I am Roman Shaffer, and I thank you for the opportunity to present our research plans on digital system dependability. I will be doing most of the talking, but Steven, as he said, will be available to take questions. Can you hear me? I'm going to talk about the goals of this research, how we hope to support and augment the current process; the motivation for performing the work, what led us to do the digital system dependability work in this way; some fundamental concepts and applicability to the regulatory assessment process. Probably the first few slides will be basic for some of you, but I'll go through them anyway to give you some background of why we're doing this the way we're doing it. An overview of the selected methodology, which is a process involving fault injection experiments, a brief discussion on specific projects that we have planned, and I say here conclusion, but that should be a summary.

The function of the Office of Research is to provide technical assistance to the various user

offices, such as NRR and NMSS to meet their respective missions, whether licensing actions, rulemaking, etcetera. We can do this in a number of ways, one of which is to supplement their staff by doing licensing reviews. Examples of this, we are performing some gas centrifuge license application reviews. Another example is we are reviewing the regulations and providing them recommendations on certain decisions they need to make. Another way RES supports the user offices is through our research products.

For the dependability research in particular, the overarching goal is to continue to support acceptability decision-making regarding digital safety systems. This means the effort will supplement and augment the current process by defining objective acceptance criteria from digital technology from a system perspective -- and there'll be more on this later -- and applying modeling tools and analysis methods that will be generically applicable to the systems that we're interested in. And this is important as we move towards a performance-based regulatory framework.

Given the complexity of digital systems, we need to understand the behavior of these systems under the influence of internal and external faults so

that we can analyze any consequent errors that might produce system failures. So if we look at the sequence that the system is operating, a fault occurs, that affects the information flow within the system, and after further processing, if there is an external adverse impact on the system that is observable we call that a failure. So it's failures, errors -- I'm sorry, faults, errors, failures. When we understand their behavior, we can characterize it and analyze digital systems for performance such as timing requirements, jitter, confirm that it does what it's supposed to do upon demand. For reliability and availability, for their failure modes, do we account for all modes, and subsystem and system safety, because interconnecting safe subsystems does not guarantee a safe system.

Another aspect of this research is to investigate if the data from this research, such as on failure modes and likelihoods, will be applicable to the probabilistic risk assessments. But this is tied more to Steven's discussion tomorrow.

Next I will discuss our motivation for undertaking this effort, such as why we need to improve our understanding of newer technologies, and also sources of faults. I'll also go over some simple

illustrations of these systems. Previous research by experts in the software and hardware fields, as well as examples of catastrophic digital system failures indicate that software can have severe defects, even after V&V. There's some work by Capers Jones who correlated the number of critical and significant errors to the number of lines of code. Some other examples are the Ariane V rocket failure, the Therac-25 deaths, the work by Koopman and Siewiorek investigating various operating systems, and the most recent example is the August 14 blackout. I believe I read something that there was a defect deep in the code that was involved with that.

There's also a greater reliance on software to perform critical functions. As you see what's being proposed to the NRC, this is quite apparent. These systems are reliant on software in safety-critical functions. There's also digital hardware components, which can have design and random defects. Some work by Avizienis and Huh studied a COTS processor and found approximately 70 defects. I think is a well known example, but I call upon it here because it ties into the work, ties into our motivation for performing this work.

Because the interaction of hardware and

software can lead to a new class of defects, we need to understand how often such defects are triggered, or how often these faults occur, and how critical they are. Do they cause death, damage to the system, or are they just an annoyance? And, given the complexity of the systems, and the significant interfacing external and within the system, what practical methods are available to determine their risk, in our case to nuclear safety? We want methods that are feasible to perform, and that can be used in our regulatory process. We don't want to take upon techniques or methods that are not timely. We'll get more into this later.

The figure represents a digital system composed of hardware and software, and various sorts of faults at different phases of the system's life. The yellow stripe outer boundary represents those development processes, design features, and operating procedures meant to prevent faults and errors from occurring. The red stripe boundary on the lower side represents those design features to handle faults and errors when they occur. In the development phase, there are requirements and specification mistakes, such as incomplete specifications. Also in this phase are mistakes in implementing the specifications. In

the operational phase there are operator mistakes, including those related to human-system interfaces, human-machine interfaces, external disturbances, such as from electromagnetic radiation, humidity, temperature, etcetera, and component defects, random failures. Internal to the system are also hardware and software faults. Now, this doesn't mean we're going to be treating hardware and software as separate components. This is just an illustration of the sources of faults.

This figure may offer a better illustration of the fault error failure sequence discussed in the earlier slide. Under certain conditions, any of these mistakes, disturbances, and/or component defects could defeat the protection mechanisms in the development and operational phases of the system's life to cause faults. For example, in the hardware/software interactions. This could potentially affect the information flow within the system, which is called an error. If after further operation there is an observable effect on the system, then that is a failure. The system is said to have failed, perhaps due to improper error handling, or occurrence of another fault.

An important aspect of assuring safety of

digital safety systems is determining the criticality and associated frequency of occurrence of faults in the hardware/software interactions. In the digital system dependability work, we will take a system point of view. Because software must execute on hardware, it is critical to understand the integrated hardware/software system, and whether or not any failures in that system lead to unsafe conditions. This is not an easy task, however, as we all know.

The system functions for fault detection and handling can be quite complex, and perhaps even the majority of system software could be devoted to fault and error handling. The methodology we have selected for the digital system dependability research can be used to exercise these functions. We can therefore analyze various classes of faults for the potential to cause unsafe conditions. The results of the research, including the data generated, could potentially be used to augment and supplement the current regulatory process as far as acceptability decision-making, and that is through the development of an objective acceptance criteria.

An overview of this methodology is the subject of later slides in the presentation.

CHAIRMAN APOSTOLAKIS: Go back please, to

your slide. What do you mean by "such software may not be exercised sufficiently." The last sentence there.

MR. SHAFFER: That means during their testing they may not test all of the diagnostic functions. They may concentrate on the safety-critical functions and not necessarily make sure that the fault diagnostics perform.

CHAIRMAN APOSTOLAKIS: And this is due to what? The fact that these are complex?

MR. SHAFFER: Could be. Could be that they're complex. Could be deadlines in the project scope, any number of things.

CHAIRMAN APOSTOLAKIS: Yes, but why does this apply to your last bullet only? That's what I'm trying to understand. You say you have much of the software is designed to handle fault detection, fault location.

MR. SHAFFER: Well, that's only --

CHAIRMAN APOSTOLAKIS: That applies to everything, right?

MR. SHAFFER: Yes, it does. This is just, we're talking -- we're concentrating on the fault detection, location, isolation, and recovery functions because the safety systems that we've approved and

we'll see in the future will have these functions built in. So not only will they concentrate on the normal safety-related functions, we'll be looking at the fault isolation.

CHAIRMAN APOSTOLAKIS: Also, if you will go on before this.

MR. SHAFFER: Sure.

CHAIRMAN APOSTOLAKIS: You may have implementation errors that you will never see until you have the right external input, right?

MR. SHAFFER: That's correct.

CHAIRMAN APOSTOLAKIS: So these are not just inputs. I mean, this is just a notional diagram, I guess.

MR. SHAFFER: That's correct, I believe I stated that.

CHAIRMAN APOSTOLAKIS: You may not --

MR. SHAFFER: That's right. It's just illustrative. It's not supposed to get all possible -- implementation mistakes are sources of errors.

CHAIRMAN APOSTOLAKIS: And now you see again my favorite subject, failure rates, and the rates, and all that. I don't think the stuff on the left has anything to do with rates. The stuff on the right does. The external disturbances, for example,

you might say have a rate of occurrence. And this is the kind of thing that I keep coming back to, that before we use Markov, or whoever, any other Russian name, you have to ask yourself what does this quantitator present? Does it model all the stuff that's useful? Requirements and specification mistakes cannot be modeled. External disturbances probably can. So that's what I mean by going to the assumptions, rather than taking the model -- component defects, I don't know. May or may not. I don't know exactly what you mean. Operator mistakes could be, could be.

So this is really the essence of it, precisely because what you have on the left there is so important for software. You see, for hardware, we don't really pay much attention to it. We have all sorts of testing and all that. But for software, this is the heart of the matter.

MR. ARNDT: We'll talk tomorrow in greater detail about what kinds of modeling we've looked at as possible ways of doing this. Although this project, you need to use some kind of models to work with, but the primary emphasis of this project is the understanding of the system, not necessarily what you do with that information in terms of what model you

use.

CHAIRMAN APOSTOLAKIS: No, but all I'm saying is that this is a good picture --

MR. ARNDT: Right.

CHAIRMAN APOSTOLAKIS: -- to put in context my earlier comments about Markov, and the rates of occurrence, and all that.

MR. ARNDT: Okay.

CHAIRMAN APOSTOLAKIS: You have strong motivation here, Roman. Several slides. You are a motivated guy.

(Laughter)

CHAIRMAN APOSTOLAKIS: That's good, though. That really -- that's nice to see that.

MR. SHAFFER: In a previous slide I mentioned that digital system faults could be triggered at system interfaces. This figure is a simple representation of a digital system where we can see various interfaces, both internal and external. We have interfaces at the inputs and outputs from and to the physical plant and humans, the human operators, which again, these include the operating environment and the HMI system.

CHAIRMAN APOSTOLAKIS: See, this is now where my comment this morning becomes more relevant.

When I asked what is the level of complexity of software being used in nuclear facilities, or in general our digital I&C. Do we really have controllers in the safety systems? And if we don't, why should I worry about this?

MR. SHAFFER: This is just an illustrative example of the systems we want to test. I could just as easily have put safety system. The safety systems take an action.

CHAIRMAN APOSTOLAKIS: My point is all this input, output, humans, and so on, the control actuators, and all that, if I don't have any systems like that safety systems in the nuclear plant right now, and as given also what was said this morning that, you know, resources are limited, why should I worry about this at all?

MR. SHAFFER: Because we do have systems like this.

CHAIRMAN APOSTOLAKIS: Safety systems?

MR. SHAFFER: Sure we do. You have the maintenance technicians, you have the operators at the control panels who are going to take action based on what these certain indications are. You're going to have actions --

CHAIRMAN APOSTOLAKIS: Digital?

MR. SHAFFER: Yes. We have safety systems in newer technologies that have gone in under 50.59. Teleperm, Common Q, and Tricon. I'm not sure Tricon's is a safety system, but they're out there.

MR. WATERMAN: This is Mike Waterman. Yes, several plants have put in digital load sequencers as part of their emergency load sequencing. I know of one plant, I believe it's the Oconee units have a digital aux feedwater system. I think that's a safety system also. And right now the systems are kind of individual modular type systems that handle one function or another, but yes, those digital systems are out there, and the progressive licensees are gearing up right now to start retrofitting.

CHAIRMAN APOSTOLAKIS: I raise this issue because if you look at the general -- and maybe it doesn't apply, but if you look at the general literature out there, those guys, you know, they look at major pieces of software, like the one that controlled the Ariane rocket and so on, and they draw some conclusions and so on. And I remember I visited one of them, I was at one of the meetings of the National Academy, the group that was preparing the National Academy report. And it was very contentious. And the main theme that one of the participants kept

coming back to was `But this doesn't apply to nuclear systems. We have very simple systems. We have very simple systems. You can't take a lesson learned from Ariane and say, well, this applies to the auxiliary feedwater system.' That's what I'm trying to do. I mean, are we taking into account the level of complexity of our digital software in our plants right now? We are not trying to solve, you know, the EuroSpace problems, or NASA's problems for that matter.

MR. KEMPER: But what we're trying to do is prepare ourselves for what's coming. Okay? You're right, what's installed in the plant right now is just a smattering of what's going to be installed in terms of digital technology in 10 years. So there's a bow wave, in my humble opinion, there's a bow wave heading towards the agency of digital upgrades that are bound to happen because of the obsolescence of analog systems. So this research will position us as a regulator to do the research that we feel is needed to estimate the dependability of these systems.

CHAIRMAN APOSTOLAKIS: This morning the issue of prioritizing the various items you have in your plan came up. Maybe if you decide to come up with some prioritization scheme in the near future,

this could be one of the considerations. What to do first, and what to do second. Because right now the plan does not prioritize, but I'm pretty sure you will have to do some prioritization at some point. And a number of criteria, of course.

MR. KEMPER: But there are priorities and a schedule timeframe, but as I say, that was developed without full buy-in of our customers, our stakeholders. And this is considered a pretty high priority project.

CHAIRMAN APOSTOLAKIS: It is.

MR. KEMPER: Right.

CHAIRMAN APOSTOLAKIS: Interesting.

MR. SHAFFER: There's also interfaces as the information flows through the embedded controller -- in this case it could be a safety system -- which is represented by the dotted line, the outer dashed line. The process variables acquired by sensors is conditioned by analog hardware, converted to digital values, and then processed by calculation and/or decision logic, which could be hardware and/or software. The flow of information continues whereby the digital values are converted to analog signals to actuate a change in the process variable being controlled. It is interesting to note that the

sensors and actuators themselves can and do have embedded controllers, such as smart sensors and digital valve actuators.

The functions shown inside the dotted line can take various hardware forms, from single integrated circuits called systems on a chip, which could be field-programmable gate arrays, and/or application-specific integrated circuits, to individual cards containing processors communicating over backlink, to widely dispersed sensors and actuators communicating over field buses or through the air via radio waves connected by network bridges, routers, or gateways over an Ethernet connection to a central controller. In our focus on safety systems, we don't have any widely dispersed safety systems. But again, this is an illustrative example of all the interfaces within these digital systems.

When we consider the role of software and its significant interaction with hardware, then the challenge of finding practical methods of assessing the safety and potential risk of these systems is apparent. From the earlier slide on the fault failure error sequence, it is possible again to get a better feeling of how fault at various points to of the system could potentially affect the information flow.

Again, an error. If the information flow has an observable effect external to the dashed line on the figure outside the embedded controller, then that is a failure. That failure then could have adverse consequences for humans or the physical plant being controlled.

I will now review some concepts and challenges of the digital system dependability effort. This figure is used to graphically illustrate the hierarchical approach to digital system design, including tolerance systems. Its purpose here is to further illustrate the complexity of these systems, and the level of effort required to analyze them. On the left side are the various layers of design and protection for the physical system and its components. As we move up the layers, our fraction increases. That means the lower layers represent physical components, such as electronics, circuits, or PN junctions, where first principles are applied. The highest layer is where system architecture is represented, such as modularity and so on, and is derived from the system specifications. The right side is the hierarchy of modeling methods and tools. Accurate modeling at higher layers could require iterating with models from the next lower level to

identify and estimate critical parameters. These models can be very complex, from millions of transistors at the circuit level to hundreds of states at the architectural level. The figure on the left side identifies possible sources of faults. Physical faults could be introduced at the lowest layer, which could then be inherited by subsequently higher levels if coverage requirements are either not met or not properly specified. Also note that new faults could be introduced at each layer, which could also be passed upwards. Those faults that defeat all layers of protection are failures.

One significant challenge is to determine the level of abstraction necessary to adequately model the hardware/software system. Though we have tools for each layer available to us, our intention is to go to the lower layers only as a necessity, because of the unique and proprietary knowledge and level of effort required to analyze at those lower levels.

CHAIRMAN APOSTOLAKIS: How is this motivation only for 3.2.2? Isn't this for everything we do in this area? This nice picture?

MR. SHAFFER: It's just laying the groundwork.

CHAIRMAN APOSTOLAKIS: It's not -- yes,

but for everything, not just for 3.2.2.

MR. SHAFFER: Yes, the point is that these systems are complex, and this was a process actually applied earlier in this work. Another illustrative example of the difficulty of building safe systems and analyzing them. That's all.

CHAIRMAN APOSTOLAKIS: Another point that would be of interest here is what does the present regulatory approach, how does it fit into this?

MR. SHAFFER: How does it fit into this?

CHAIRMAN APOSTOLAKIS: Yes.

MR. SHAFFER: In our current approach?

CHAIRMAN APOSTOLAKIS: Yes.

MR. SHAFFER: Well, as you've heard earlier, we focus mostly on the software development lifecycle, but then there's also --

CHAIRMAN APOSTOLAKIS: So what is that? I mean, we're covering all these architectural level, algorithmic level, functional level. I mean, we do that?

MR. SHAFFER: No.

CHAIRMAN APOSTOLAKIS: No. Yes? Yes or no? You said yes? You want to come to the microphone? Identify yourself, please.

MR. CHIRAMAL: I'm Matt Chiramal from NRR,

and this is -- we look at every level of this. But maybe talk to the BT that he's talking about, but we look at all the levels, architecture, algorithmic, functional, logic, circuit level. These are parts of the review.

CHAIRMAN APOSTOLAKIS: On the left, you mean? Every level on the left? Although the right is really modeling.

MR. CHIRAMAL: On the right is when they start designing it completely. At this point, the SER is on the platforms.

CHAIRMAN APOSTOLAKIS: Okay, thanks. Let's go on.

MR. SHAFFER: For safe operation, a digital system must have the capability to detect a large percentage of faults. When a fault is detected, the system will perform appropriate action to prevent transition to an unsafe state or condition. In the dependability community, the parameter for measuring how well a system prevents unsafe conditions after detecting a fault is fault coverage, or simply coverage.

Coverage is defined as a conditional probability that the system correctly handles a fault, given that a fault occurs. Note that there are --

CHAIRMAN APOSTOLAKIS: That's not the same way Mr. Li defined it earlier, is it?

MR. SHAFFER: These are different projects, different methods, different areas of focus.

CHAIRMAN APOSTOLAKIS: Different terminologies.

MR. SHAFFER: He's talking about test coverage. We're talking about coverage from the fault tolerant dependability community.

CHAIRMAN APOSTOLAKIS: But his test coverage was not a condition of probability, was it?

MR. SHAFFER: You'll have to talk to him about that. It's not my project.

CHAIRMAN APOSTOLAKIS: No, it's not.

MR. ARNDT: The effort he was talking about was a software testing concept of how much of a particular set of code was covered during a particular kind of testing. This is a different concept which just happens to use the same -- similar terminology.

CHAIRMAN APOSTOLAKIS: So C_D is the conditional probability that a fault exists and we don't detect it?

MR. SHAFFER: That's correct. Now, C_D is the probability given that there's a fault that your fault detection functions detect it. Given that there

is a fault, it's the probability that the fault detection circuit will detect that fault. A failure would be 1-C_D, and that would be a coverage failure.

Note that there are different types of coverage. For simplicity, the term "coverage" will be used to reference a system's coverage requirements. Coverage requirements are application-specific. A failsafe system would require high fault detection coverage in order to shut down to a safe state, whereas a highly reliable system would require fault recovery mechanisms to restore the system to a known good state after detecting a fault. Note recovery requires fault detection, fault location, fault isolation, and fault recovery. Coverage is an important concept, but it is a difficult parameter to estimate.

CHAIRMAN APOSTOLAKIS: I don't understand the probability C_I. Why is there a probability that the fault would be isolated? Can you give me an example?

MR. SHAFFER: Again, it has to do with the function in the software code or the hardware.

CHAIRMAN APOSTOLAKIS: If I know where the fault is, and say I know the redundancy of the system, shouldn't I know with certainty whether this is

isolated or not? Why do I have a probability that it will be isolated?

MR. SHAFFER: Because it may not perform its function all the time. I mean, there's --

CHAIRMAN APOSTOLAKIS: I don't understand why that would be the case.

MR. SHAFFER: Why it would be the case? Because circuits fail, hardware fails. There's just certain failures in a system where the fault isolation circuit may not work.

MR. ARNDT: Take for example if you have a fault tolerant system, either software fault tolerant or hardware fault tolerant, that compares the output of a sub-routine, or compares the output of a processor. If for some reason the system has a fault that affects both of those, then you're not isolating the fault. There's some probability that --

CHAIRMAN APOSTOLAKIS: You are really unlucky, in other words. Not only is there a fault --

MR. ARNDT: Well, that depends on our architecture.

CHAIRMAN APOSTOLAKIS: -- what you have built into the system to protect you against it also fails.

MR. ARNDT: Right.

CHAIRMAN APOSTOLAKIS: And then fault recovery would be the conditional probability that all these terrible things have happened, but still I recover somehow?

MR. SHAFFER: And your system handles the fault correctly, in this case yes. That it recovers correctly. If any of those fail, then it's considered a coverage failure, and you end up in an unsafe condition.

CHAIRMAN APOSTOLAKIS: I guess you're going to give us some examples of this.

MR. SHAFFER: Okay. Watchdog timer detects a fault, resets the system, it's a fault recovery mechanism. For fault recovery you can go to your checkpoints when you detect a fault. To recover from that, you can either go back in time to a known good state, or you could go forward to repair the system and find -- starting out in an error state, you eventually transition to a good state, a normal operations state. Interrupt service routine. That can be considered a forward recovery mechanism in software.

CHAIRMAN APOSTOLAKIS: Anyway, keep going.

MR. SHAFFER: A number of researchers have developed methods to assess the reliability of digital

systems, Jeff Voas, Jacob Abraham, Kang Shin, Ravi Iyer, Koopman and Siewiorek, Barry Johnson, and Jay Lala among others. Two current issues for the NRC regarding digital safety-related systems are understanding the behavior of digital safety systems, and understanding the risk of digital safety systems. This project is focused on the former, with the hope to provide relevant data for the latter under a different project, which Steven will discuss maybe during this presentation, if you have questions, or tomorrow.

The digital system dependability research will undertake several case studies to attempt to estimate the coverage of qualified digital systems. These systems all have built-in diagnostics. Because these systems were designed to different requirements, not only will the research give us more insight into the safety of the systems, but also the research will allow us to apply the method to diverse platforms for different reactor applications. The objective is to determine if their built-in fault tolerant protection mechanisms function as expected, or fail under certain conditions, and if they do fail, what are the consequences. We want to determine the criticality of the failures.

Longer term we want to know if the selected methodology provides credible results, for example, under a peer review. That's an important component of this research. We need to bring in experts from diverse fields and have them review our work. And is it practicable, that is the method has measurable benefits to the current regulatory process for the level of effort it requires.

The presentation will now turn to an overview of the selected methodology shown in the figure. More detailed information is available in technical reports generated during a cooperative agreement with the University of Virginia. There is a report associated with each of those blocks. The research will build upon the UVA effort by applying the process to digital safety systems. These projects will be discussed in more detail later. UVA originally developed this method for designing safety-critical systems as they have been involved in about 20 different system design projects. They've actually built fault tolerant systems with this methodology. NRC intends to apply the process to assessing several safety-critical systems as case studies. The process is based on an effective technique for characterizing system behavior under faulty conditions called fault

injection. By injecting corrupted signals either onto hardware pins or into software instruction sequences, it is possible to determine how the system will react. The fault injection experiments will be used to estimate critical model parameters necessary for solving the derived analytical model, which is the first block there.

The process starts with determining reliability and/or safety requirements, and confidence levels, and deriving an analytical model, perhaps using Markov models, Petri nets, or even fault trees. Because this is a quantitative approach, system information generated from certain qualitative analyses, such as design reviews, hazards analyses, etcetera, will be used when developing the analytical model. The statistical models for estimating the critical model parameters, in our case coverage, using input from the fault injection experiments. The statistical model determines the number of fault injection experiments required to meet the confidence intervals. The remainder of the process essentially determines the types of faults to inject based on expected operational profiles in order to measure internal operating parameters of the system for later analysis.

There are several issues we need to address before this technique can be practically applied to the NRC's process. For example, the fault space of the system could be extremely large, thus requiring a large number of fault injection experiments to obtain a statistically significant set, which could be impractical given the length of time required for each test.

CHAIRMAN APOSTOLAKIS: It's not just the faults. It's also the external inputs. You inject the fault, then you have a whole space of external inputs.

MR. SHAFFER: That's correct.

CHAIRMAN APOSTOLAKIS: Two big spaces, actually, isn't it?

MR. SHAFFER: Well, the idea, again coverage is a conditional probability given that a fault exists. It doesn't care the source of the fault, whether it's an operator action, whether it's a random hardware failure. The faults represent conditions of the system as a result of a fault. The fault represents conditions of the system under certain adverse consequences.

CHAIRMAN APOSTOLAKIS: Given one fault --

MR. SHAFFER: Which could represent

anything.

CHAIRMAN APOSTOLAKIS: -- there's a lot of space of inputs.

MR. SHAFFER: That's correct. And it could represent inputs, hardware failures, whatever.

CHAIRMAN APOSTOLAKIS: What is the typical number of faults in these applications people have produced?

MR. SHAFFER: In this process that UVA has applied, they have injected over 100,000 faults in one case.

CHAIRMAN APOSTOLAKIS: There is an intelligent way for defining those faults?

MR. SHAFFER: Yes, at the lower blocks there, 4, 5, 6, and 7, that's where the detailed knowledge of the system is required. Further compounding the problem of the large fault set is the issue of no response faults. Assuming a tractable sample set of experiments could be found, it is possible that many of the faults selected will not result in any noticeable effect on a system. These are called no response faults. These are essentially latent errors that have not caused any noticeable effect for the duration of the experiment. Other issues related to practicality include actual

construction of the test harness, how we can actually perform the fault injection experiments, and test automation. How do we, as we perform a test, and we get a response that may lock up the system, there has to be some way to automatically reset the system. Because if you need an operator there to reset every time, the total test time could be intractable, given the number of experiments that have to be performed.

The digital system dependability research will allow confirmation that the fault injection process we have selected addresses these issues sufficiently enough so that it can be applied to digital systems of interest to the NRC. We want to effectively determine how safety systems behave under faulted conditions. Such information could potentially be used to augment and supplement the current process for reviewing license applications, and that direct testing of qualified systems in approved configurations could lead to realistically conservative licensing decisions, based on both deterministic and probabilistic criteria.

An illustration of what we plan to do is in this figure. We will have the capability to model both the hardware, the software, and its interfaces. Because we will have physical access to the systems,

we will have the actual code. However, I do not discount the potential need for alternative software models. These are very complex systems. The hardware model is based on simple fetch-execute computer architecture. Again, we'll have physical access to the system, and we have a generic processor model which is one of the blocks in the figure on process.

MEMBER WHITE: Excuse me, Roman.

MR. SHAFFER: Sure.

MEMBER WHITE: Are you going to also handle common failures? In other words, multiple faults?

MR. SHAFFER: We will handle multiple faults. Now, whether they're common mode, we believe we'll be able to use the results of this to address that issue. Whether we will actually be able to define what a common mode failure is, particularly a software common mode failure, I am not sure we'll be able to do that because then we would need more than one channel.

MEMBER WHITE: Okay, thanks.

CHAIRMAN APOSTOLAKIS: But when you select the faults, in general you don't have common cause failures in mind?

MR. SHAFFER: We're going to have a huge

fault space. To really do these common mode failures, we would probably need more than one channel, and inject faults.

CHAIRMAN APOSTOLAKIS: But you could do it?

MR. SHAFFER: In principle, yes. But the scope of our work is a single channel.

CHAIRMAN APOSTOLAKIS: And the fault can be a software problem or a hardware problem, failure?

MR. SHAFFER: That's correct. We're going to mess with --

CHAIRMAN APOSTOLAKIS: And it's always one?

MR. SHAFFER: Well, what we've found on previous work is if we inject a single fault, then sometimes we see multiple corruptions, multiple corruptions being faults at multiple locations in the system. In fact, up to five. Those are a very small percentage, but we've seen that.

CHAIRMAN APOSTOLAKIS: Okay.

MR. SHAFFER: The generic processor model, which we will discuss in a moment, will enable us to determine the types of faults to inject. However, long-term, if we could develop a process that was not dependent on having the hardware available and would

still allow in-depth analysis, that would be ideal. There is potential to develop a simulation model of the hardware configuration, and use that for simulation-based fault injection. And that will be discussed later as well.

As I said earlier, we have modeling tools that allow us to go to the gate level, so that is always an option. But we're always looking for efficiencies in our processes. If we can stay at a relatively high level of abstraction, that sort of releases us from having actual hardware, but then we become dependent on the vendors and the engineers, those who have real knowledge of the system.

Now we're going to discuss each block one by one. This is just an overview. The analytical safety model provides the mathematical framework for calculating reliability and/or safety estimates. It's simply a high-level representation of the faulty behavior of the system under analysis. Several suitable analytical models from the literature include Markov models, Petri nets, fault trees, and variations, colored Petri nets, dynamic fault trees, etcetera. Critical fault parameters may include failure rates, repair rates, fault detection latencies, and fault coverage. This is the most

difficult to estimate, but that's the parameter we want to estimate.

CHAIRMAN APOSTOLAKIS: It seems to me you will have here the problems we discussed earlier, namely whenever in real life, or even in your testing processes, you find faults, you probably fix them. So.

MR. SHAFFER: Yes, during a design process you would --

CHAIRMAN APOSTOLAKIS: Even in whatever process. I can't imagine that you find, you know, faults and you just leave them there. Maybe one or two you say I don't care, but in general you go and correct the problem. So now, you know, the parameters you want, again, is the -- are the statistics collected applicable. This is a really tough problem, you know. By the way, this is not unique to you. NASA had that huge problem with the shuttle. Every time they find a problem they fix it, and sometimes the fix costs half a million dollars. And here comes now the risk analyst saying `Oh, there were five failures' and the guy goes bananas. I spent half a million eliminate this problem, and you're telling me that it's still a failure. So I don't know. I mean, this estimation of remaining faults from things that

I have found and I have fixed is something that I don't think we know how to handle as a community.

MR. SHAFFER: Can I give you a little background maybe?

CHAIRMAN APOSTOLAKIS: Yes, you can give me background.

MR. SHAFFER: UVA developed this initially as a way to design fault tolerant systems. It would work in parallel between hardware and software, where you would catch the faults early. We happened upon this at a later time, and determined that it may be useful to an assessment process. Our intention is to obtain certain qualitative analyses where we may already have certain information available to us, and from there determine what the design safety requirements were. And from there then we could establish, you know.

CHAIRMAN APOSTOLAKIS: Are you coming back to Steve's argument of earlier today that, you know, no matter what the numbers are, at the end I have gained a hell of a lot of insights to the system by doing this. And I'm 100 percent with you.

MR. SHAFFER: That's right.

CHAIRMAN APOSTOLAKIS: I mean really, if you inject 100,000 faults and you find what's going

on, I mean more power to you. But when you start calculating lambdas like Dr. Johnson did here several months ago, then I get cold, to the point of freezing sometimes.

MR. SHAFFER: Yes, well --

CHAIRMAN APOSTOLAKIS: I just don't think you can do that. And I'm willing to listen. I mean, I'm dying to find an argument that says this is the right thing to do. I don't see it. I haven't seen it. And it's not your problem. It's not your problem. Don't take it personally.

MR. SHAFFER: No, I don't.

CHAIRMAN APOSTOLAKIS: Nobody knows how to do that, including me.

MR. SHAFFER: I think that --

CHAIRMAN APOSTOLAKIS: We're on the same boat. Sergio, you're smiling. Do you know anybody who can do it?

MEMBER GUARRO: No. That's surprising, that expression of modesty, that's all.

(Laughter)

MR. ARNDT: I can say that's uncharacteristic that he should.

CHAIRMAN APOSTOLAKIS: I mean, do you disagree with anything I just said? No. No. And I'm

telling you, we had that problem in NASA PRA of the shuttle. And as you know, how political that is now, right? Re-launching the shuttle after the accident and so on. It was a real problem. Here you have a guy who says `I just spent a quarter of a million dollars fixing this problem, and you're telling me you're going to consider it a failure and do the calculations as if nothing happened?' What do you say to that? So they came up with a methodology for discounting failures. So this was not one failure, this was 0.65 of a failure, you know, that kind of a thing. And you appreciate now what kind of issues come out of that. But it's a real issue. It is a real issue. And I think we have that here too.

MR. ARNDT: We do.

CHAIRMAN APOSTOLAKIS: There are two arguments. One is can you really ignore some failure that happened because you think you fixed it, and second, by trying to fix it, have you introduced additional problems. So anyway, as far as the analysis of the structure of the software/hardware I have no problem with that. I mean, all this method clearly gives you good insights. But when we go to numerical estimates, now I don't know. Okay. So let's go on. Unless you disagree with what I said.

MR. SHAFFER: I think there's --

CHAIRMAN APOSTOLAKIS: Look, I'm trying to learn here. It's not -- but somebody has to be the bad guy here.

MR. SHAFFER: I don't think you're being a bad guy at all. In the -- for whatever that's worth.

MR. KEMPER: Let me step into this for just a second. Maybe I shouldn't, but certainly the intent, our desire is to come up with some way of substantiating the reliability of this system. That's what we desire the licensees to be able to demonstrate to us.

CHAIRMAN APOSTOLAKIS: I'm with you. What I'm saying here is that these are big issues. They are not just your problem. And as a community, we don't know how to attack them, and the sooner all of us agree to that, and then start from there, the better off we'll all be. Because I've seen a lot of applications where people take existing models from reliability theory and they force them onto software because, you know, it's the standard thing. You know, I've lost my keys and I'm looking around the lamp because that's where the light is. So.

MR. KEMPER: Well, there may not be an

obvious solution at this point, but we're certainly going to continue to pursue that.

CHAIRMAN APOSTOLAKIS: Your use of the word "obvious" was very unfortunate.

(Laughter)

MR. KEMPER: Doesn't that mean that a solution cannot be achieved, right? I tell my folks all the time, the world was flat for a long time until we proved that.

CHAIRMAN APOSTOLAKIS: These days it's triangular.

MR. KEMPER: That's right.

CHAIRMAN APOSTOLAKIS: Okay, Roman. Sorry for the interruption.

MR. SHAFFER: No problem.

CHAIRMAN APOSTOLAKIS: I'm not really sorry.

(Laughter)

MR. SHAFFER: Okay. The statistical model is used to estimate critical model parameters in the analytical model.

CHAIRMAN APOSTOLAKIS: I would skip this.

MR. SHAFFER: Why?

CHAIRMAN APOSTOLAKIS: We've discussed this enough. Keep going.

MR. SHAFFER: Well, this is an important component.

CHAIRMAN APOSTOLAKIS: I know, and now you're -- okay.

MR. SHAFFER: Well, we use the statistical model to determine how many fault injection experiments we do. And that's a critical component of --

CHAIRMAN APOSTOLAKIS: What did you say?

MR. SHAFFER: We use the statistical model to estimate, or to determine the number of fault injection experiments to perform. Okay?

CHAIRMAN APOSTOLAKIS: I'd like to see that. There may be some value to it. Yes, I agree.

MR. SHAFFER: So we have single --

CHAIRMAN APOSTOLAKIS: But you haven't fixed anything. Yes, good. That's fine.

MR. SHAFFER: The statistical model is also used to determine -- I'm sorry. Okay. The statistical model is used to determine the number of fault injection experiments, but also that in turn affects which fault injection technique we'll use of the four. We'll discuss these later.

CHAIRMAN APOSTOLAKIS: I'll tell you what. The statistical model I'm sure has value, but what

would have more value as far as I'm concerned is to see some intelligent way of selecting the faults based on the anticipated use of the system.

MR. SHAFFER: That's where the novelty of this approach.

CHAIRMAN APOSTOLAKIS: That's where I would really love to see how they do that. You know, pretty soon before you realize it you have to understand all the accident conditions you might have in the plan, right? Because these are safety systems, so they have to respond and control, if you will, say accident situations. And my God, you're getting into accident space. I don't know. Dr. Kress, do you think we understand all that?

MEMBER KRESS: I think you do have to get into accident space.

CHAIRMAN APOSTOLAKIS: In which case it's a huge space.

MEMBER KRESS: It's a huge space.

CHAIRMAN APOSTOLAKIS: And I'd like to know whether there are any intelligent ways, or semi-intelligent ways of selecting where to put the fault. Not just the number of faults, but also where.

MR. SHAFFER: Well, in this process they apply those algorithms.

CHAIRMAN APOSTOLAKIS: I'd like to see that. I mean, I'm sure Dr. Johnson, does he have anything? Because he did them for trains. I don't know, but maybe you guys could do it.

MR. ARNDT: Yes. One of the outputs of this particular project will be looking at how do you apply those kind of methodologies that have been used in other --

CHAIRMAN APOSTOLAKIS: Failure and in nuclear.

MR. ARNDT: Right.

CHAIRMAN APOSTOLAKIS: Okay, great.

MR. SHAFFER: UVA developed a behavior level model of a generic processor, a basic fetch-execute cycle. It was applied to a design project in Europe, and was certified by TUV Germany. The generic processor fault model is used to generate the fault space for the system, where the fault space is defined by location, time, and value. Location is where the fault occurs within the system under analysis. Time is the time of occurrence and duration of permanent or one instruction cycle. Value is a defined corruption of the correct entity called a mask. Any accessible registers and memory locations can be corrupted. Detailed fault models have been derived from the

literature for register file and register selection faults, program counter faults, control unit, instruction decode logic faults, data address control bus faults, and arithmetic logic units. This generic model was validated by simulation, and augmented by refining the masks. And then it was applied to several COTS processors, two Motorola and an AMD. For digital system dependability research, the generic model will be applied to the processors and the systems under test, and then an appropriate fault space generated, which again could be very large. Therefore, certain techniques to reduce the number of fault injection experiments to a tractable number will have to be used.

Before performing the fault injection experiments, however, the system is placed into context by determining appropriate operational profiles. If it's an RPS, we'll have to define a proper operational profile, if it's load sequencer, etcetera. These should be representative of the system under various modes of operation and configuration, since various configurations may invoke different hardware and software functions. To get a good understanding of the system's behavior under faulted conditions, a sufficient number of

combinations should be analyzed.

The operational profile is divided into four phases, a startup phase, where the system is allowed to reach a stable state, no faults are injected due to the short time interval relative to the operational time. It's statistically insignificant. The second phase is a system light workload where there are no faults from the simulated external environment, and thus only a reduced set of software and hardware functions are running in the background, such as diagnostics. The third phase is a system heavy workload where significant interaction with the simulated external environment to exercise as much of the system's resources as possible. And the fourth phase is a short no activity phase so that outputs can stabilize to determine externally observable effects due to the fault injection. Then you determine if the system failed. This sequence will thoroughly exercise the system and allow us to observe its behavior under the influence of both transient and permanent faults.

After determining the appropriate set of operational profiles, the experimental setup will simulate the selected operational environment under fault-free conditions. Data will be collected on the

system under test, such as instruction sequences of observable state data, system buses, etcetera. Note that the fault diagnostic functions are also monitored, and information is collected. This data here is the fault-free execution trace. Equipment used includes logic analyzers, bus analyzers, in-circuit emulators, and software debuggers. So we're going to get a lot of information.

The set of injected faults and the analysis of the fault injection experiments are dependent on the fault-free execution trace. For example, when a fault is injected into the system, data is again collected on the system's response and compared to the fault-free trace. Therefore, the fault-free execution traces should have as much detail as possible to ensure accurate identification of covered, uncovered, and no response faults.

One significant challenge with fault injection is that the fault space can be quite large, making it unfeasible to test the entire fault space. A reduced set of faults is then randomly selected from the fault space. Recall that this statistical model determines the number of fault injection experiments that must be performed to satisfy the confidence intervals. Another challenge, however, is that not

all faults injected cause an observable effect. Therefore, the initial statistically significant subset does not provide enough data to estimate the critical model parameter. Further, the no response faults are the worst case result as far as time of testing. Fault injection tests that yield no response faults require the longest amount of time as the system response is compared to the fault-free execution trace. So you're waiting for a response that doesn't come during the duration of the test. So they're just long tests.

To overcome problems posed by no response faults, a technique to collapse the fault list by eliminating no response faults is applied. This is based on work by Benso, Guthoff, Smith, et al, and Iyer, Ravi Iyer, et al. However, there still leaves the issue of a large set of tests to inject as determined by the statistical model. For systems with high coverage requirements, the number of required fault injection experiments may be quite large. The concept of fault equivalence may be applied to reduce the number of experiments. This is essentially a variance reduction technique. The algorithm seeks to identify sets of faults that have an identical effect on the system, even though each fault in the set is

distinct. These sets of equivalent faults are called equivalence classes. Only one fault from each equivalence class needs to be injected to determine the effect of all faults in that class.

The earlier algorithm was received with some criticisms, so they refined it. The assumption is the faults are uniformly distributed in the fault space, therefore they have equal probability of occurrence. They randomly sample a number of faults, and they determine the number of equivalent classes from those faults. Since with assumption one there's no bias in the coverage estimates since the faults in the equivalence classes are also random.

Again, the effectiveness depends on how much information can be derived from the execution trace. In a real world example, UVA applied the process to an interlocking control system, which is a failsafe application of 10 years of operation, 150 locations throughout the country, 30,000 lines of assembly code, had a time requirement of 200 millisecond response time, and 80 percent of the code was devoted to diagnostics. They injected over 100,000 permanent faults. And using this fault expansion technique, that approximated about 9.5 billion faults. They evaluated about 1,900 transient

faults. And that was expanded to about 136 million faults. Clearly that offers us some advantages. Again, as I said earlier, the work will undergo a peer review, so there will be time for scrutiny of the results. And getting back to Steven's point, just doing the fault injection experiments, having a set of faults that we know will get a response, and determining the system's response will give us a large amount of information. I believe that'll be useful to the safety reviewers as well as the PRA.

CHAIRMAN APOSTOLAKIS: So this is now from the Virginia work, this kind of diagram?

MR. SHAFFER: Yes. It is, actually. In fact, most of these slides are.

CHAIRMAN APOSTOLAKIS: So you plan to adapt it to nuclear applications?

MR. SHAFFER: We do indeed.

CHAIRMAN APOSTOLAKIS: You did already? Or you will?

MR. SHAFFER: We're undergoing a feedwater control system assessment now. And we will apply this to safety-related systems.

CHAIRMAN APOSTOLAKIS: Because I was thinking, as I said earlier, maybe you need something there addressing the issue of environments, accident

environments. Somewhere in there, you know, you have to have that.

MR. SHAFFER: Again --

CHAIRMAN APOSTOLAKIS: Given the fault, what are the possible inputs to the software. If I have a small LOCA, a large LOCA, if I have this, if I have that. Those are different inputs.

MR. SHAFFER: Right, but again, coverage is a conditional probability that doesn't care about the source of the faults. Given a fault, does the system detect it.

CHAIRMAN APOSTOLAKIS: Right. But what if you miss a whole class of inputs because you never considered a medium LOCA? Then you can't find the conditional probability because you missed a lot of possible inputs, given the fault. That's what I'm saying. Given the fault, you may have a whole space of possible inputs depending on the accident.

MR. KEMPER: Roman, I believe back on Slide 19, is that where you? I assume that you were addressing that when you said light loads versus heavy loads for the operational profiles?

MR. SHAFFER: Yes, that's part of it. Yes.

MR. KEMPER: Okay.

CHAIRMAN APOSTOLAKIS: But what I'm saying is that you need to show it explicitly in those figures.

MR. SHAFFER: Okay, well there's a way we can -- when we determine the fault space, it's possible for us to trace backwards to what the external inputs would be, or could be. I mean, given that --

CHAIRMAN APOSTOLAKIS: Well, I'm not saying you can't do it, Roman. All I'm saying is that, you know --

MR. ARNDT: Be sure to do it.

CHAIRMAN APOSTOLAKIS: Yes. I didn't say you can't do it.

MEMBER GUARRO: I think you mentioned the assumption that the faults are uniformly distributed in the fault space.

MR. SHAFFER: That's correct.

MEMBER GUARRO: And is that a valid assumption?

MR. ARNDT: Well, that is not a necessary assumption. It just happens to be the going in assumption. You can go in and do a parametric study to look at what the distribution is, and/or what effects it may have depending upon your assumed input

states.

MEMBER GUARRO: Yes, because I'm thinking of an analogy. You tell me if it's out of context. But I'm thinking of the difference between a pure Monte Carlo sampling and a Latin Hypercube sampling, in which you're worried about, you know, characterizing details that are rare events. And so now you go there more often than you should under a theoretical assumption or uniformity. I think probably something like this, my intuition tells me that may apply. I may be wrong.

MR. ARNDT: It's a similar concept, although not exactly the same thing. And the point is well taken. The Virginia work did do some work on statistics of the extreme to look at this as part of applying this to a nuclear example, and George's point that these are rare events in many cases, and it's difficult to characterize them. You have to go back and carefully, as Roman was saying earlier, if you have a particular fault, you can go backwards and look at the input state that's associated with that. So what you need to do is you do the experiment, then you start relaxing assumptions, and look at does the uniform distribution as opposed to a different kind of distribution have an issue. Is the fault space you're

using characteristic of an appropriate operational profile and input characterization. Those are parts of QA'ing the process to make sure it is applicable to a nuclear example.

CHAIRMAN APOSTOLAKIS: Roman keeps telling us that we are looking for a conditional probability giving the fault. In a nuclear application, it's not inconceivable that you will have a number of conditional probabilities, namely given this fault, and given I have a small LOCA, here is the conditional probability of it. Given the same fault, but given that I have a large LOCA, maybe I have another conditional probability. So it's a double condition, in other words. It doesn't sound too far-fetched to me. I mean, different accidents create different conditions.

MR. ARNDT: Right. And you can characterize those conditions, those accident conditions if you will, as input parameters. For a trip circuit you have low pressure.

CHAIRMAN APOSTOLAKIS: Sure, but I would like to know these conditional probabilities. And if you just tell me given this fault the conditional probability of failure is 10^{-3}, maybe you're not giving me the whole story.

MR. ARNDT: Yes, that goes back to what is the operational profile and what's the fault space.

CHAIRMAN APOSTOLAKIS: Okay, okay. Let's go on. It's getting late in the day, and Steve threatens us with two presentations here.

MR. ARNDT: Yes. They're both short.

CHAIRMAN APOSTOLAKIS: Yes I know about that. Risk assessment, short.

MR. SHAFFER: We have several fault injection methods available to us. I won't spend too much time on these. We have hardware-based fault injection, which is essentially where we augment the system with additional hardware so we can perform the whole fault injection experiments. We have software-based fault injection, and that's where we develop a -- we modify, interrupt service routine to inject changes in the software operation. A simulation-based fault injection is where we have a complete simulation model of the system. There is commercial software available called SIMEX where they provide complete models of certain microprocessors. We've considered doing that. And then the final approach is the hybrid approach, which is some combination. It's possible we could do a simulation of the processor interface to he hardware prototype and perform a series of fault

injections that way. But this will all have to be determined during the project.

Now, there are advantages and disadvantages, but I don't need to go through those. We'll get into the research projects. Over the past few years, we've done a Digital Feedwater Control System assessment, and it's continuing under the cooperative agreement with OSU. The second project is the Digital System Dependability Performance project, which will kick off in the end of FY05. And this is a multi-year effort. This is the project where we're going to evaluate a number of systems. We believe there's great benefit to all parties involved here, but mostly to us because we get a better assurance of safety of these systems. We'll know how they fail, and we'll be able to incorporate that into our process. Right now we have three platforms that we've generically approved. This work doesn't propose to redo all that. We want to look at these in their -- as close to site-specific implementations as we can.

Future effort will explore other dependability metrics, such as maintainability, confidentiality, and integrity. That's under the security work, which from my understanding we'll come before you again later and discuss those, which are

also my projects.

The Digital System Dependability Performance project for the highest probability of success will require work with vendors and licensees. We'll have access to the systems, but we're also going to need access to their systems designers, engineers. They're the ones with the knowledge of the malicious faults. Those are the faults they know that if they occur, an unsafe condition could happen. I'm not saying that these systems are unsafe in any way, but there are certain conditions that if they happen, if the protections are defeated, could lead to adverse consequences. During the work, we'll perform the fault injection testing following the process described earlier. And we estimate about 12 months per system evaluation. It's actually platform.

CHAIRMAN APOSTOLAKIS: How does this dependability work different from risk assessment? Isn't this part of what you have to do to do a risk assessment?

MR. ARNDT: To do a risk assessment you need to, as you know, understand the ways the system can fail.

CHAIRMAN APOSTOLAKIS: And this helps me do that.

MR. ARNDT: This helps us do this. This is going to be one possible input to the supporting analysis necessary to build failure models for a PRA. But the reason we've got it as a separate broken-out project is, first of all, just the way systems fail is not in and of itself just a reliability issue, it's an understanding the system better, as well as, if you go back up one slide, you can use these methods to do other things, like integrity of the system, to look at things like the security-type issues as well. You can look at other dependability metrics other than failure rate.

MR. SHAFFER: I think I should state that as I've been talking there was an implicit assumption that these safety systems we've approved have unsafe failures, unsafe faults. It could very well be that we don't find anything. We don't know. I don't want to say ahead of time that they do.

CHAIRMAN APOSTOLAKIS: Well, if you find anything it will be a small number.

MR. SHAFFER: This is true. But the idea is that we know, and that's where everyone benefits. It's all about assurance for us. And if it's assurance for us, the licensees have assurance.

I do say conclusion, but I mean to say

summary. The digital system dependability research will augment and supplement the current regulatory process by characterizing significant hardware, software, and interface errors, including system interface errors that could prevent safety system action or cause initiating events which could undo the challenge-mitigating systems, understanding potential new failure modes and the criteria for detecting these failure modes prior to failure of plant safety functions, identifying or developing methods and data that enable the NRC to establish the risk importance aspects of digital safety systems, Steven's project, and modeling of digital systems that could be used to support probabilistic risk assessments. And that's all.

CHAIRMAN APOSTOLAKIS: Comments or questions from the members or the consultants?

MEMBER WHITE: I have one question on your generic process fault model. You were talking about time, and you said that would include the fault injection time, and the duration, and the duration would be either one cycle or permanent, I think. Do you think you might eventually look at fault durations that are intermittent, you know, just for several cycles, then off? You understand what I mean?

MR. SHAFFER: Yes. That's a slightly more difficult problem, but yes. We could reach that point.

CHAIRMAN APOSTOLAKIS: Anything else? Well, thank you very much. Now, Steve, why don't we go ahead with your self-test methods.

MR. ARNDT: Okay.

CHAIRMAN APOSTOLAKIS: I think that's next, right?

MR. ARNDT: Yes. Just for scheduling points of view, this will be a relatively short presentation, and then I've got about a half hour, 45-minute presentation on the overview of the risk program.

CHAIRMAN APOSTOLAKIS: Maybe we'll take a short break between the two?

MR. ARNDT: Yes. That would be good.

MR. SHAFFER: Did you say self-test methods?

MR. ARNDT: Yes.

CHAIRMAN APOSTOLAKIS: So let's finish this because with the next one, we start the whole issue of risk assessment.

MR. ARNDT: Well, this was originally intended to be a fairly short presentation because we

haven't really done a lot of background work on this. This is just something new that we're going to be starting, and we wanted to give you some general overview. When Mike Waterman gave this presentation this morning when talking about the discussion of the comments on the research program he discussed a lot of this, so some of this will be redundant, so I'll go through this relatively quickly.

As we talked about this morning, this program is under the Software Quality Assurance program. It need not necessarily be there. It could have been under the emerging technology part of the program, or the systems aspect program. The reason we put it here as opposed to some other place was a lot of these issues are software issues. Not all self-testing is software. Some of it's hardware. But this just seemed like the easiest place to put it.

As we discussed this morning, self-testing methods can be hardware or software tests that are done on a continuous basis to improve the system available. They're designed into the system to improve the availability or functionality of the system. This is distinguished from a subject that we have in another part of the plan that talks about system diagnostics. That talks about is the system as

a whole working, or system issues associated with evaluation of calibration and things like that. This is aspects of the system that are specifically designed to improve the hardware/software, the digital part of the system.

One of the issues associated with this over the years was the overhead associated with these. That's pretty much gone away. Even with real-time safety-critical systems, the power of these systems from a computational standpoint has significantly reduced the overhead issues associated with that. The performance issues are different. The issues associated with is the system going to have an issue associated with too much crammed into a cycle time, or locking the diagnostic system up, or having a fault in the diagnostic system affecting the performance of the overall system. Those issues still exist. It's just, the point of the bullet is the fact that because the overhead is not such a big deal, these systems are more commonly used.

And these can be very, very simple kinds of things, like checking to make sure that the system has executed all of its programs in the allotted time, various kinds of very simple self-checks. It can be inversion programming kinds of things to determine

whether or not two different versions of the code came up with the same answer, and then going forward. There's a number of different kinds of things that we're talking about.

One of the big issues, as NRR talked about earlier when they presented, as we talked about earlier in several programs is the complexity issue. The idea of these systems is to improve the availability by making sure the thing doesn't fail when it doesn't have to. But the problem is you're adding additional complexity in the overall system as you add more and more self-checking type applications. So the real issue here is we want to understand, one, is there a tradeoff between how much complexity you add and the failure modes associated with the added complexity and the actual system itself. The other thing is are there systems or types of self-checking that are preferred as opposed to not preferred. An analogy would be an effort that we did a few years ago on safe programming language applications. We did a study on what was the preferred methodologies for coding. The idea behind that project was to give NRR a potential list of things that are likely to be good coding practices, and things that might not be so good. When you see them in a review, you need to

spend more time looking at them.

As we talked about this morning, there's a limited amount of time you can spend in a review. And you have significant time resources associated with that. So the idea of this project, the outcome, is to provide additional information to the regulatory review staff on aspects of self-testing that they might want to look at more closely. What does the experience tell us? What does the theory tell us associated with what's the best way to do these, and where might there be some problems?

I've gone through a lot of these in the overview. The issue is what effects, if any, might this have on system performance, what adverse effect may it have, what are the most acceptable testing methodologies versus the lease acceptable testing methodologies, and what is the theoretically best or most acceptable amount of self-testing. So the project is basically going to focus on those kinds of aspects, as well as what operational history has told us. Mike Waterman this morning gave you two examples of systems in nuclear applications that failed because of self-testing issues, not because of the actual systems that they were designed to -- the functional aspects of the system. There's been a lot of cases in

clearly the software part of it, and there's been examples in other areas where because they didn't think through some of the ways systems -- the diagnostics could fail, they put them into application, and they had faults because the complexities associated with the self-test got the best of them. So. How much self-testing is enough? How much is too much? What kind is appropriate is really what we're trying to look for, both from a best practices operational experience, and theoretical standpoint.

So that's what this project's about. We haven't kicked it off yet. As Mike mentioned before, we'll probably have a lot of interactions -- we intend to have a lot of interactions with our NRR colleagues associated with this. We've discussed this with them once already on what aspects of this they think is most appropriate. And we'll go through the process of --

CHAIRMAN APOSTOLAKIS: Have you decided who's going to do this?

MR. ARNDT: No. We have not decided.

CHAIRMAN APOSTOLAKIS: Okay. Thanks Steve. You say your next presentation is a 45-minute presentation?

MR. ARNDT: Half hour, 45 minutes. Depends on how many questions we get.

CHAIRMAN APOSTOLAKIS: Okay. So let's recess till 4:20.

MEMBER KRESS: Let's go ahead and do it. We don't need a recess. Let's go ahead and do it.

CHAIRMAN APOSTOLAKIS: No, let's break for awhile.

MR. ARNDT: Let's break.

CHAIRMAN APOSTOLAKIS: Okay, 15 minutes.

(Whereupon, the foregoing matter went off the record at 4:05 p.m. and went back on the record at 4:24 p.m.).

CHAIRMAN APOSTOLAKIS: Okay, Mr. Arndt. Risk assessment. You're speechless.

MR. ARNDT: Absolutely. I'm in awe by your greatness.

MEMBER KRESS: Bow down.

MR. ARNDT: The purpose of this presentation, like the overview presentation that Bill and I gave earlier in the day is to give some general background on the overall risk assessment program, get some general ideas on why we think we should be doing it, why we think it's important, and the structure of the overall program. Tomorrow we will go into the

individual programs and some intermediate results associated with them.

As we all know, the NRC has a PRA policy statement which encourages the use of PRA to the extent supported by the state-of-the-art and data. One of the big issues that is central to this is what is the state-of-the-art. Do we have sufficient information and techniques to be able to do this kind of work? And it really gets to, and I'm going to talk about this a little more in a couple of slides, the fact that there's two issues here. The issue that we've been primarily focusing on is is the state-of-the-art such that we an inform the regulatory process in approving and evaluating digital systems for applications based on risk-informed information. The other issue, of course, is that all the rest of the risk-informed applications are based on a complete PRA. And of course, as the licensees put more and more digital systems into the plant, a general PRA that doesn't model digital systems and their interactions is less complete. So we have both those issues as potential outcomes and issues associated with this.

So the research is oriented toward improving the NRC's knowledge and providing consistent

processes for regulating digital systems. So what we want to do is all the kinds of things you want to do when you develop risk models. We're going to gather and understand the data, assess the modeling methods that might be used, what is adequate, understand the systems that need to be modeled, and what level of detail. This is one of the big issues. Like any other modeling application, there may be some models that can be modeled fairly simplistically, and there may be some systems that you have to model at a much greater level of detail simply because of the complexity of the system, and/or how they interface with other systems. We have to develop and test methods. Now we don't necessarily have to develop them ourselves, but we have to understand what the modeling capabilities are, what the limitations are, and whether or not we can live with those limitations. And then we have to develop regulatory acceptance criteria. This is the point we made earlier. By acceptance criteria, what we mean is those aspects of digital system analysis in reliability space that are particular issues for digital systems. So for example, regulatory acceptance criteria might be a version of the 1.74 series specific to digital systems that highlights those additional issues that you want

to look at in digital system analysis.

As we've talked about extensively, licensees are replacing analog systems with digital systems. In some cases, these are fairly sporadic occasional type issues in non-safety systems. In other cases, like the Oconee case, they're looking at doing a very complete digital system replacement of a large number of safety systems, trip systems, SFAS systems, and things like that.

Some of the current deterministic licensing criteria are challenges. The one that has been most in the news recently is BTP-19, which is the staff guidance on diversity and defense-in-depth. One of the challenges associated with this is how that analysis has to be done. The industry has expressed interest in using risk-informed ideas as an alternative method for meeting some of these more challenging issues, like diversity an defense-in-depth. And I'll talk about that briefly later. So there is some interest in using risk information, or risk perspectives in the current licensing framework. So the real issue is what are the limitations of digital system reliability models, and can they be used, can the be expanded, can they be used in a limiting kind of a thing, or some certain aspects, or

do we need to do a full modification of PRAs.

The real issues are not easy ones, as Professor Apostolakis mentioned. In addition, even if we got a risk-informed application, our current methods and data within the agency to do an independent assessment aren't up to par either. So if we get an application either in the forms of a topical report, and we have one for review, or of an actual application based upon a risk-informed application, we currently don't have methods available to us to do an independent assessment.

CHAIRMAN APOSTOLAKIS: Now, can you -- maybe you covered it, I don't know. Let's go back. I'm intrigued by the sub-bullet that says that some of the current licensing criteria are difficult to meet. Can you give an example or two?

MR. ARNDT: The example that is used is the diversity and defense requirement. The way the diversity and defense requirement, BTP-19, is written, you have to do an analysis of what would happen in the case of a common mode software failure. The recommended analysis associated with that, and someone correct me if I don't get this quite right, makes certain assumptions that basically says if you have a software failure, you have to assume a large part of

your systems fail, and then go through all your design basis accidents and determine that even with this software failure, you can withstand in Part 100 space the design basis accidents. Now, some of that's not very difficult to do because you have auxiliary backup systems which are not safety grade. You have operator actions. You go over and punch out the system and things like that. But there are some accidents that that becomes a real challenge for. Large-break LOCA is the one that comes to mind, and that's primarily because of the timing issues associated with it. So because that is a deterministic analysis making certain what most people would call very conservative assumptions, you have some challenges in meeting that.

Now, the alternative is you put it in a diverse backup system in addition to your digital systems. Now, obviously if you believe your digital systems are of high quality and reliable in the first place, you don't want to have that added burden associated with them. But when I say some current licensing criteria are difficult to meet, it means there are certain criteria that if you take them at their base, they're believed by many in the industry to be overly conservative and force you to make design tradeoffs they would prefer not to make. Did I get

that roughly correctly?

MR. TOROK: May I comment?

CHAIRMAN APOSTOLAKIS: Identify yourself, please.

MR. TOROK: My name is Ray Torok. I'm from EPRI, and I'm the project manager on the industry guideline on this subject. And all I was going to add to what Steve said there was that in a case like the large-break LOCA, obviously it's a low probability kind of event, but also what you find when you look at it in PRA space is that the probability of failure of the system is dominated not by the INC in the system, but by the large rotating machinery, so that even if you add a diverse backup like Steve's talking about, from a risk standpoint it doesn't help in terms of core damage frequency and so on. And it does add complexity that may actually increase the likelihood of a problem.

MR. ARNDT: So, in any case, the point is there are reasons that the industry is interested in some form of risk-informing some of our regulations because of these kinds of issues. How exactly that's done --

CHAIRMAN APOSTOLAKIS: So ultimately you would like to be able to use Regulatory Guide 1.174?

That's really what you would like to do?

MR. ARNDT: The industry has, that's the approach it would like.

CHAIRMAN APOSTOLAKIS: Well, you do. I mean, the Commission's policy is to be risk-informed, right?

MR. ARNDT: Yes.

CHAIRMAN APOSTOLAKIS: All of us.

MR. ARNDT: And we'll go into that more. In the June 2004 ACRS letter, Professor Apostolakis also in his added comments recommended that in this particular area, databases containing software-induced failures should be reviewed, and their contributions should be used, the information we gained from that. And he also recommended available methods for assessing reliability systems that are software-driven should be reviewed critically. And this is a bit of a paraphrase, but I believe that's generally the idea.

MEMBER KRESS: You realize, of course, that the reason these are added comments is the rest of the ACRS rejected them.

MR. ARNDT: I understand. I was there when --

CHAIRMAN APOSTOLAKIS: So the next slide will not do either of these, right?

(Laughter)

MEMBER KRESS: Remember, we attempt to accommodate all perspectives.

CHAIRMAN APOSTOLAKIS: Thank you very much.

MR. ARNDT: Both of these are very appropriate comments. In deriving what we're trying to accomplish, we need to understand -- and they both go to the issues associated with the PRA policy statement. What we want to do is understand what the state-of-the-art is and what the state-of-the-data is, and what we want to do is build on that in our Research Program Plan. So the point of highlighting these here is it goes back to my first slide. What we're trying to do is understand the state-of-the-art, build on the state-of-the-art, and try and get to where we need to be, which is both the policy in terms of how we're going to interface with the licensees, and also our own internal methodologies.

So the research program is designed to use the available information in data to understand the capabilities, as I said on the last slide. The big issues here are to look at what's going on and use the most promising methods, or at least try to use the most promising methods and investigate them. We

really have two issues here. We need to understand what is and is not possible, and what are the limitations of the modeling effort. We need to do that for two reasons. One, to direct our research in the right way for internal analysis techniques, but also to help us support development of regulatory guidance. So when EPRI or one of the licensees comes in with an application, we understand what the limitations are so we can ask better questions. So what we want to do is work on that.

So as part of our program, we're going to look at, and develop, and integrate new methods. And "new methods" is probably too strong of a word. It's new methods to the NRC. We also want to pilot these things using both traditional methods and dynamic methods where appropriate. We want to benchmark the capabilities of different methodologies. One of the biggest issues, of course, in any new methodology is you need some benchmarks. How well did these work in specific applications. So as we talk about what is exactly in our program, one of the things we want to do is for certain applications, for certain kinds of systems, we want to benchmark the different kinds of methodologies that have been proposed, and understand based on both the theoretic aspects and the benchmarks

what the limitations are. So as part of that, we can get out guidance for regulatory applications.

And my EPRI colleagues have chided me already on this first bullet. EPRI has proposed a methodology. The biggest issue associated with that, which is not a sub-bullet, but please pencil it in, is the fact that their methodology relies on the measures that are designed into the system to enhance its reliability. Things like fault tolerant behavior, and things like that. They want to take credit for how these systems are designed. They rely also on the issue that Ray just brought up, that a lot of the systems, total systems, not just the digital systems, have aspects associated with the failures of the big spinning parts. So their methodology looks at understanding the system from a total system perspective, particularly the bounding assumptions associated with the reliability of the digital system compared to the system it is controlling, or it's actuating.

CHAIRMAN APOSTOLAKIS: Has EPRI submitted this report for formal review by the NRC?

MR. ARNDT: It's been submitted. It's under what is known as acceptance review consideration right now. As part of review of topical reports, we

look at it and say, all right, do we want to review it. And this has a number of different issues associated with it, associated with resources, and do we think it's an acceptance methodology, and do we want to review it or not. So right this particular second we're trying to determine if we're going to review it, and what the schedule's going to be. Yes?

MR. TOROK: May I offer a couple more comments?

MR. ARNDT: Sure.

MR. TOROK: This is Ray Torok from EPRI again. And yes, I just wanted to offer a couple of clarifications there. The first bullet says it's a method for incorporating digital systems into current generation PRAs. And I would characterize maybe a little differently in that what we were trying to do was apply risk insights to defense-in-depth and diversity evaluations for digital upgrades. Now, that does lead you to addressing the issue of modeling digital systems in PRA. They're obviously related. And what we do is we use estimated failure probabilities for the digital equipment to get it into the same playing field as the other hardware in the system that the digital licensee happens to be. So in that sense it's a qualitative approach, really, where

you apply engineering judgment, and in some cases standards, like the one that's mentioned, the IEC standard. But it really comes back to engineering judgment at some point.

Now, what Steve mentioned about these defensive measures things is very important for both determination of susceptibility, where you may be susceptible to the common cause failure, and for estimating failure probability of the digital equipment. We go back to looking at these defensive measures that are built into the digital system. And that's really important because it gets you beyond just looking at the process. Because what you really want to know is what the real system behaviors are, and make your decisions based on that. Because there are large uncertainties in the digital equipment failure probabilities, we address that now with uncertainties, which means that if the NRC research work comes up with better ways to determine those probabilities of failure, they're certainly applicable within the framework. So I see that as all fitting together in a nice way. Thanks.

MR. ARNDT: Thank you, Ray. Our research is focused a little bit differently. We're focusing more on the, if you will, the fundamentals of the

reliability modeling, understanding what kind of models might be appropriate, developing both traditional and dynamic methods, and seeing whether or not they're applicable, if they work, where they work, investigating model acceptability, and doing some benchmarks. So we're going at it from a slightly different perspective. As Ray mentioned, hopefully our framework will be sufficiently broad that we can include what they're doing, and they're hopefully going to do the same thing. So the issue really is we're attacking it from slightly different perspectives, but the objective is to have a methodology where we can include risk insights into the regulatory process.

This is a historical graph, and I'll only spend about a minute on it. This is what I presented last March when we talked about this. I found a better way of doing it, so I'm just putting it up here to remind you. The concept is there are certain aspects of this that we're investigating. We're trying to understand PRAs and digital systems in them as our final product. To do that, you need to understand the digital system itself. You have to understand the hardware, the software, and the supporting analysis that provides you the failure

modes, and things like that. As we've tried to develop this program a little bit better, we've come up with a somewhat I hope better way of looking at all the different aspects of our program, which is on Slide 9.

What we're really doing from a project standpoint within the program is trying to accomplish certain things. If you look at the left-hand side of your screen, one of the aspects that's very important in both choosing what kind of models you do, as well as supporting the models, is understanding what the failure data is. Another aspect is reviewing the current reliability modeling methodologies, and coming up with ideas on what might work best, choosing the candidates for possible inclusion. Those both tie into the development of approaches for modeling the systems.

That center box there is really what we were talking about this morning, and early this afternoon, supporting analysis. You need to understand how this system works in one way or the other to be able to characterize it in some kind of model, be it a fault tree model, be it a dynamic flow graph model, be it any kind of model. You need to understand how the system works, and not just the

visual system, but how it works with the process that it's interfacing with. And you can do this in a number of different ways. You can use hazard analysis, you can use failure modes and effects analysis, you can use software testing, or fault injection, or a number of other methodologies to understand how the system works. Those we don't include in the digital system reliability program because those are outside the program, but they're feeding into it. Those are the things that we need to understand to develop the reliability models.

On the left-hand side is traditional method, fault tree/event tree modeling methodologies, and on the right-hand side is dynamic methodologies. One of the big issues, as Professor Apostolakis has highlighted, and a lot of other people have highlighted, is there's a lot of ways you could potentially do this. A lot of ways that people do it. And there's a lot of argument. Well, is this appropriate. Is that appropriate. Can you do this. There is no consensus in the community. We need to, one, understand what the limitations of the various models are, and also we need to understand for our internal needs what is the best way to do this. One of the biggest issues is when you model these things,

what are the limitations. And are you willing to live with the limitations for that particular application. The whole point of developing a PRA quality standard is saying, all right, in some applications you can live with a less sophisticated model. In some applications you can't live with a less sophisticated model. You need a greater amount of details, or a better understanding of things. To write a regulatory position on that, be it a Reg Guide 1.17x, or be it into the quality standard, or whatever, you need to have an appreciation of that.

To do that, what we've done in our program is specifically had two different sets of researchers looking at it from two different aspects, and trying to independently assess whether or not this is possible, and what the limitations are in particular cases. As we develop methodologies to do that, then we're going to also develop benchmarks. Right now we're looking at two benchmarks that have certain aspects associated with them. One would be a control system, probably an aux feedwater system. It has less, quote unquote "safety significance" but it has a lot more process interactions. The other would probably be a reactor protection system trip signal. It has a lot more safety significance, but from a

modeling standpoint it's much simpler. We want to get some benchmarks that give us the biggest bang for the buck, we learn the most from doing the analysis. And the idea would be to use two or three different methodologies, both traditional fault tree/event tree methodologies, and maybe some of the dynamic methodologies, and understand both from a modeling perspective and an understanding of how hard or easy it is to actually do these kinds of models. Based on that, we will then decide how to, or if we should update NRC tools and data to provide independent assessments.

Now, I've spent a fair amount of time talking about the graph, and the next three or four slides basically are redundant to what I've just said, but I'll go through them quickly anyway. But this is really the concept behind what we're trying to do. And what we'll talk about tomorrow is particular pieces and parts of that.

CHAIRMAN APOSTOLAKIS: I'm not disagreeing with anything you said, but my -- the thrust of my comments in the letter that you cited, but also other people's comments, is that in this particular case of software, we shouldn't just jump into Markov models or whatever. We should really question the basic

assumptions behind it, precisely because the evidence is that most of the problems come from specification errors, requirements, you know, design type errors. So it's really a different way of thinking about reliability models. And it's very easy to just say, oh okay, well I'll use a Markov in a discrete state, and move from here to there. What does this lambda 223 mean? What are the random events that you're assuming are occurring, and you know, at a constant rate? So this is really the critical review that I was talking about. And I think it's important to do that, and I assure you we can do it.

MR. ARNDT: Right. And we specifically -- and I agree. And there's two issues associated with that. One, you have to do as good a job of reviewing possible strengths and limitations in the various models as you can before you start spending money to do development. And we think we've done a pretty good job, and you're going to hear some of that tomorrow. The other issue is at some point you have to start doing a little bit more detailed analysis and modeling to understand the limitations. Can you choose any particular methodology, dynamic flow graph methodology for example? What are the limitations in terms of practicality? Can you get enough data? Can you get

the prime implements? Is there a methodology that can be used so you have both the understanding of the limitations as well as understanding the practicality.

CHAIRMAN APOSTOLAKIS: But the most fundamental thing is the theoretical basis. You know, you can have a practical method that is not theoretically sound, you're in trouble. That doesn't mean that you go with the best theoretical method. Practicality comes in, there is no question about it, but the theoretical basis I think is very important.

MR. ARNDT: It is.

CHAIRMAN APOSTOLAKIS: And there is literature on these issues. I don't know if you guys have found it. In the past people have argued back and forth.

MR. ARNDT: Yes, we've done a fairly sophisticated review of a lot of the literature, including the paper you referenced in your additional comments, among others. Both the development of a theoretical -- or the set of assumptions we're going to choose to use, I should say. It's not so much a theoretical argument, but it's a choice of what arguments we're going to choose to use, as well as an evaluation of what seems most promising is something that we're specifically working on.

CHAIRMAN APOSTOLAKIS: All right.

MR. ARNDT: Let me reiterate a couple of things as I go through these other slides. Like I said, most of it I've already talked to you about. The outcome is to really understand what systems need to be modeled, what level of detail they need to be modeled, what kind of accuracy are we talking about, what uncertainty, if you will, are we talking about. Developing the capability to independently verify these systems, and developing acceptance criteria. What is we want out of the licensee application. So as I mentioned before, we're specifically looking at several different methodologies. We've got two different research teams specifically so we don't miss anything, so we look at it from several different aspects.

This is the part of the project that's looking at the data. We're going to have some more discussion on it tomorrow so I won't dwell on it. But as part of this, we're looking at what's out there, what can be used, what more information do we need. One of the biggest problems is most of the digital failure databases don't have enough information in them to support reliability calculations directly.

CHAIRMAN APOSTOLAKIS: But this is all

nuclear you mean? When you say failure data, nuclear failure data? Or you're looking at other industries?

MR. ARNDT: We're looking outside as well. But I'll give you one example, the LER database, which is used for a lot of different things. The problems associated with that are numerous. It will give you some digital system failures, but in many cases it doesn't give you a sufficient level of detail to characterize it in one way or the other.

CHAIRMAN APOSTOLAKIS: But we also don't have extensive experience with these things, do we?

MR. ARNDT: We have less than great experience in many areas. We don't have time between failures, we don't have number of systems deployed, and issues like that to get basically the denominator in the equation. So there's a lot of issues associated with it, but we want to use as much data as we can, if nothing else to inform the process, but also to develop these kinds of databases that are going to be needed.

The purpose of part of our research is really to understand what is out there, what are the advantages and disadvantages. And I've talked about this fairly significantly. The issues associated, what the risk-important characteristics are, what are

the things you have to include in the model is something that's very important. Basically the methodologies, which ones are the most appropriate. Can you use one kind of methodology, or another kind of methodology, and if you use one kind of methodology, what limitations do you have to place on your results?

So in summary, the research is designed to solve basically the issues that we have. And we've also designed it as a broad-based program looking at a number of different potentially viable options. And one of the things we really, really, really want is to have a proactive interactive relationship with the subcommittee on these issues. Because this is a controversial issue, we're trying to build in peer review wherever it makes sense. To some extent you can't peer review everything or all you do is spend time making presentations like this. But wherever it makes sense, we want to get interaction with the technical community, be it papers, and conferences, and journal articles. We want to get interactions with the licensee community. We're planning to have a workshop probably summer, late summer, fall, to talk about some of the aspects of the regulatory issues that we're looking at. We've had some external peer

reviews on some of our research products. We'd like to have that same kind of rapport with the committee. And what we're really looking for is how can we do things better, where are things that might prove more promising, and issues like that. Things that we may not have considered, or you think we might consider more, those kinds of interactions are something we would appreciate.

Like I say, we're going to go into some of the details much more extensively tomorrow, both in terms of planning for each of the blocks that we had up here, as well as some of the preliminary results we've had in a couple of the areas.

CHAIRMAN APOSTOLAKIS: Okay. That's it?

MR. ARNDT: That's it for my overview.

MR. KEMPER: That's all we have to present today. So if you'd like to continue on tomorrow we could do that I guess. Or we could continue on this afternoon if you prefer.

CHAIRMAN APOSTOLAKIS: Any comments, questions, from our people around the table? No? Jim, no? Shall we go around the table you think, or should we do it tomorrow afternoon? Jim and Sergio will send us a written report.

MEMBER KRESS: So we can wait till

tomorrow.

CHAIRMAN APOSTOLAKIS: Wait until tomorrow I think. Okay. Any comments from the audience?

MR. WATERMAN: Professor Apostolakis, I just wanted to add one thing that Jim White pointed out during the break was in all of these projects in the research plan, you'll notice the last product is a training curricula for whatever the product might be such that not only do we have, for example, a review procedure, but we also intend to incorporate into our contract some form of curricula development so that when we deliver that product to our supported offices they also get training on how to use that product in a consistent manner, which is just absolutely critical. Instead of just dropping something on somebody's desk and saying `Now, go use' we really want to emphasize that all of these things need some form of curricula developed so that as new staff come on down the road they can be sent off to be trained on how to use those products, and so we can build up our infrastructure so that people like Paul Loeser aren't just on their own. It's unfortunate that we have to use GS-14s and GS-15s to do a lot of the grunt work that you can take a kid straight out of school to do, but right now we're kind of stuck with you need an

expert to review every aspect of the system because that's all you've got. And a training curricula is designed to help resolve some of that by providing some of that expertise to some of our junior staff members to bring them along. Thank you.

CHAIRMAN APOSTOLAKIS: Okay. So, thank you very much Steve. And this first day is over. We'll reconvene tomorrow at 8:30. No? Well, this subject at 1:00. 8:30 we have another meeting. Right? Okay. Thank you.

(Whereupon, the foregoing matter was concluded at 5:02 p.m.).