Thank you to both reviewers and the editors for your time and feedback. We have made a number of improvements to the paper to address the comments. Our responses provided below in blue.

Reviewer: 1  
  
Recommendation: Accept If Certain Minor Revisions Are Made  
  
Comments:  
The paper describes a set of tools that enable the formal analysis of a system specified using Architecture Analysis and Design Language (AADL). The tools incorporate, among other things, mechanisms for analyzing the cybersecurity of the system and developing cybersecurity requirements. The tools extend to the analysis of the designed and implemented system as well, so it's possible to assess whether the identified requirements are met by the implementation. The tools have been applied in the context of a design for an unmanned helicopter and in particular to modifying the design so that it incorporated a wireless tablet that could monitor data about nearby air traffic. By basing the architecture on the verified sel4 platform and applying various tools to assure first that the system specification meets its requirements and then that the implementation of the specific can be assured to conform to it, the project aims at the holy grail for high assurance systems, verification down to the code level.  
  
The paper is reasonably clear and should be of interest to a broad audience. It reports a promising result -- that a team of development engineers, most of whom had no previous experience with formal methods, was able to use the tools effectively to meet the requirements of the demonstration. The parts of the paper describing the methodology and tools are quite detailed and might be improved by raising the level of the discussion a bit. Perhaps an overview describing at a high level the assurance argument for the entire system would eliminate the need for some of the detail. The account of the application, on the other hand, might benefit from additional detail, if available, about the training required (or not required) of the development time, the time consumed in accomplishing the tasks involved, and so on.

We have eliminated some detail throughout the BriefCASE Workflow section to raise the level of discussion as suggested. We have also added a new Figure 2 providing an overview of the tool architecture showing how the individual tools interact. Hopefully these changes will help readers not to get lost.

The figures, with the exception of Figure 5, are barely legible in the pdf and will be illegible in print. Figure 4 is especially bad in this respect. They need to be redone, and possibly recast, so they actually add value to the paper. Also note that the description of Figure 1 asserts that the Planner module is white, but in fact, it's pink (and the reason for that is not explained).

Figure 1 (Cyber-resilient software architecture) has been updated to make the component names larger and eliminate port names which were too small to read. The caption and associated text have been corrected regarding colors.

The previous Figure 3 (Awas Information Flow) has been eliminated since it was too busy and hard to read anyway.

Figure 4 has been revised to zoom in on one subtree of the assurance case for illustration, making the text there readable.

Additional Questions:  
1. How relevant is this manuscript to the readers of this periodical? Please explain your rating in the Detailed Comments section.: Very Relevant  
  
2. To what extent is this manuscript relevant to readers around the world?: The manuscript is of interest to readers throughout the world  
  
3. What is the most appropriate forum for the publication of this manuscript?: IEEE Magazine (general interest explanatory article with technical contributions)  
  
1. Does the manuscript contain title, abstract, and/or keywords?: Yes  
  
2. Are the title, abstract, and keywords appropriate? Please elaborate in the Detailed Comments section.: Yes  
  
3. Does the manuscript contain sufficient and appropriate references (maximum 15-unless the article is a survey or tutorial in scope)? Please elaborate in the Detailed Comments section.: References are sufficient and appropriate  
  
4. How would you rate the organization of the manuscript? Please elaborate in the Detailed Comments section.: Could be improved  
  
5. Is the length of the manuscript appropriate for the topic? Please elaborate in the Detailed Comments section.: Satisfactory  
  
6. Please rate and comment on the readability of this manuscript in the Detailed Comments section.: Readable - but requires some effort to understand  
  
1. Please summarize what you view as the key point(s) of the manuscript and the importance of the content to the readers of this periodical.: The paper describes a set of tools that enable the formal analysis of a system specified using Architecture Analysis and Design Language (AADL). The tools incorporate, among other things, mechanisms for analyzing the cybersecurity of the system and developing cybersecurity requirements. The tools extend to the analysis of the designed and implemented system as well, so it's possible to assess whether the identified requirements are met by the implementation. The tools have been applied in the context of a design for an unmanned helicopter and in particular to modifying the design so that it incorporated a wireless tablet that could monitor data about nearby air traffic. By basing the architecture on the verified sel4 platform and applying various tools to assure first that the system specification meets its requirements and then that the implementation of the specific can be assured to conform to it, the project aims at the holy grail for high assurance systems, verification down to the code level.  
  
2. Is the manuscript technically sound? Please explain your answer in the Detailed Comments section.: Appears to be - but didn't check completely  
  
3. What do you see as this manuscript's contribution to the literature in this field?: It provides evidence that formal methods tools can be applied effectively by developers without special backgrounds and explains the specific tools and techniques used.  
  
4. What do you see as the strongest aspect of this manuscript?: The example it provides.  
  
5. What do you see as the weakest aspect of this manuscript?: Some of the description is perhaps overly detailed. See notes below. Most of the figures are likely to be illegible.

Addressed above.  
  
7. Please rate and comment on the timeliness and long term interest of this manuscript to S&P readers in the Detailed Comments section. Select all that apply.: Topic and content are likely to be of growing interest to S&P readers over the next 12 months  
  
Please rate the manuscript. Explain your choice in the Detailed Comments section.: Good  
  
  
Reviewer: 2  
  
Recommendation: Accept If Certain Minor Revisions Are Made  
  
Comments:  
Overall, this was a really excellent paper that I enjoyed reading. The paper presents some great research which genuinely extends the limits, scale, and applications of formal methods technologies, fully meeting the brief for this special issue.  
  
As stated above, I was really excited to read about such a range of formal methods tools, techniques, technologies, and verified components all used in one project, all contributing to a combined assurance case in a rigorous manner. This is a really important demonstration of how various tools can work together, and how outputs from different tools can be combined meaningfully to support the highest levels of assurance.  
  
This paper presents low 'tech readiness level' research rather than finished products, so I think in reality it's unlikely that many avionics scenarios would be able to use the whole suite in such a comprehensive way for now. In spite of that, the presentation and combination of this array of tools is a fantastic piece of work, and strongly demonstrates the use of these formal methods tools at real-world, high-assurance scale.  
  
Major comments:  
---  
  
I found it difficult to get the structure of the tool approach on first pass: this is a big toolset with lots of moving parts, and it required lots of effort on the part of the reader to construct a model of what each of the individual things do and how they work together. I would like to see some 'high level', introductory summary paragraph and/or graphic of how all the moving parts relate.

We have added a new Figure 2 providing an overview of the tool architecture showing how the individual tools interact. Tools to be discussed in the BriefCASE Workflow section are highlighted in green. We hope these changes will guide the discussion and help readers not to get lost.

Some more transparency is needed over a couple of areas, specifically real-time requirements and programming language choices.  
Although the BriefCASE toolset and workflow is not only targeted at avionics, the paper has a very strong avionics emphasis. In this light, much more consideration and discussion in the paper needed about domain-specific (avionics) requirements needed to address concerns.  
For example, there is no meaningful discussion about real-time / WCET safety requirements in the avionics domain; RTOS is alluded to, and the 'Real-time scheduling' section starts to address this, but does not provide sufficient discussion about how these real-time safety critical requirements are met (or not) within the technologies used. This needs to be discussed more clearly (and should not be difficult to do so!), or I feel it will not be taken seriously by those in avionics.

We agree that the real-time aspects of the design are important, but it was not actually a focus of our work on the CASE program. Our main objective was to address cyber-resiliency in the design process. The applications we chose for demonstration had timing constraints related to resource consumption and throughput objectives, but they were not truly hard real-time (like control loops). Consequently, some of our design choices (like using CakeML for high-assurance components) might have been different if we had different program objectives or demonstration targets.

To make the real-time discussion a bit more complete we have added some additional detail regarding the need for verifying WCET of tasks in additional to scheduling guarantees in the 2nd paragraph of the Real-Time Scheduling section (page 8). However, this is not currently part of the BriefCASE tools. This whole topic of support for real-time tasks will be better addressed in the MCS version of seL4 (once its verification is complete).

Similarly, more discussion is needed about the apparent tension between the use of functional programming languages such as CakeML and the real-time/WCET requirements of the very conservative avionics domain (which would normally require manual memory management). The paper even mentions the JVM at one point (re: HAMR, p.9): I acknowledge that this output is not a core part of the toolset, but use of this in safety-related avionics context is seemingly implausible.

HAMR’s code generation support for JVM is only provided for prototyping, as indicated in the 1st paragraph of the HAMR section (page 8). Additional discussion in this section of the Slang and JVM support for prototype development has been eliminated to avoid confusion. The topic is addressed better in other publications.

As far as the decision to use CakeML, this was driven by the assurance goals of the project. Certainly, garbage collection has to be taken into account, but for our very structured use of the language it did not prove to be a problem. Some acknowledgement of these factors has been added to the 2nd paragraph of the High Assurance Component Synthesis section (page 6).

The paper would also benefit from a brief discussion about satisfaction of various safety standards. For example, there are supplements to DO-178C that cover MBSE, formal methods, and tool qualification: how does this approach relate to this or other relevant standards?

A short discussion on how key standards are supported by the assurance case approach has been added to the beginning of the Assurance Case section (page 10).

The paper should also make clear (Aircraft Application, p.12) whether the "vision processing module" that seL4 is running on is one for which seL4 is fully verified or not, i.e. is the VPM system's platform one for which the proofs are complete? I don't see it as a substantial problem if not, but transparency and clear descriptions over precisely what each element of verified technology (or analysis tool) does or does not buy the assurance methodology as a whole is essential to maintain trust in any approach.  
  
This is a good point. The ODROID-XU4 (ARM Cortex A7) that was used for the UAV Surveillance example system included fully verified version of seL4. While the Xilinx ZCU-104 (ARM Cortex A53) and the Collins hardware equivalent used for the helicopter demonstration system used the same seL4 code base, it was not fully verified for that platform. An explanation has been added to the end of the Motivation section (page 3) and the next to last paragraph of the Aircraft Application section (page 12).

Minor comments:  
---  
  
Figure 4 is blurry when zoomed in so can't see much of the assurance case  
p.2, line 54, col 1 - "Proofs about models are meaningless unless..." - I appreciate the sentiment but it's a bit hyperbolic

Updated to be less hyperbolic (“Proofs about models have less value…”)

p.4, "BriefCASE work flow": Where does your list of cybersecurity vulnerabilities come from? What is the threat model, and where does this come from?

Cyber-resiliency requirements are generated by the two requirements tools that we support (GearCASE and DCRYPPS). These were developed by other researcher teams who have different approaches to identifying vulnerabilities. Generally speaking, both have internal representations of what constitutes a vulnerability and overlap largely with the MITRE CVE list.

p.4, "Requirements": At what layer or level are these vulnerabilities? How complete do you claim this approach is?

The vulnerabilities address by the requirements tools focus on the architectural level, the system boundaries, and interconnections between components, rather than on low-level binary or hardware vulnerabilities.

p.4, Section "Cyber Transforms" - Is there only one possible transform at each stage? If not what happens if there's more than one at any given stage? How did you choose this list of possible transforms, and what level of completeness do you believe you achieve with these transforms?

The library of transforms is not complete and is mainly intended to illustrate the approach. It can definitely be extended to address other cyber requirements. In some cases, the vulnerability identified and the associated requirement clearly suggest a unique transform. In other cases, multiple requirements are addressed by a single transform. Our current list was chosen based on the needs of our example systems.

p.5, Source/citation needed for Lustre

Citation added.

p.6, line 43, col 1 - missing full stop after "contiguity types"

This section has been deleted in the final version.

p.6, line 54, col 2 - "verified using Coq" - doesn't explain how

The attestation managers themselves are written using Coq to verify attestation protocols and protocol interpreters, then those models are synthesized to CakeML. This is noted at the end of the 2nd paragraph of the Remote Attestation section (page 7).

p.7, para 1: Does this attestation process only happen once, and then allow all messages (unauthenticated/unsigned) from that source? If so, how does the UAV know that message after this point aren't spoofed, replayed, or from another source? I assume there's some cryptographic protocol used here (creating an encrypted and/or authenticated channel), but this is not described.

Attestation can be requested periodically if warranted by the anticipated threats against the remote system. This has been added to the 3rd paragraph of the Remote Attestation section (page 7). Periodic re-attestation was implemented in the helicopter demonstration system.

Messages between the tablet and the avionics equipment are encrypted. See the last paragraph on page 11 in the Aircraft Application section.   
  
Additional Questions:  
1. How relevant is this manuscript to the readers of this periodical? Please explain your rating in the Detailed Comments section.: Very Relevant  
  
2. To what extent is this manuscript relevant to readers around the world?: The manuscript is of interest to readers throughout the world  
  
3. What is the most appropriate forum for the publication of this manuscript?: IEEE Magazine (general interest explanatory article with technical contributions)  
  
1. Does the manuscript contain title, abstract, and/or keywords?: Yes  
  
2. Are the title, abstract, and keywords appropriate? Please elaborate in the Detailed Comments section.: Yes  
  
3. Does the manuscript contain sufficient and appropriate references (maximum 15-unless the article is a survey or tutorial in scope)? Please elaborate in the Detailed Comments section.: References are sufficient and appropriate  
  
4. How would you rate the organization of the manuscript? Please elaborate in the Detailed Comments section.: Satisfactory  
  
5. Is the length of the manuscript appropriate for the topic? Please elaborate in the Detailed Comments section.: Satisfactory  
  
6. Please rate and comment on the readability of this manuscript in the Detailed Comments section.: Easy to read  
  
1. Please summarize what you view as the key point(s) of the manuscript and the importance of the content to the readers of this periodical.: This paper presents and describes a comprehensive new approach to use of formal methods for large-scale, systems engineering assurance. This uses a 'model-based systems engineering' approach (and environment), integrating formal methods throughout, and across a range of different layers of the engineering design stack. This has the benefit of being able to rigorously address and mitigate a range of cyber-security issues at the design stage, rather than much later as is more usually the case. This paper demonstrates new combinations and use-cases for tooling, and shows the utility and possible scale of use-cases for formal methods-based analysis tools and techniques.  
  
2. Is the manuscript technically sound? Please explain your answer in the Detailed Comments section.: Yes  
  
3. What do you see as this manuscript's contribution to the literature in this field?: The combination of such an array of tools and verified components/technologies in this domain, into a genuinely comprehensive suite for high assurance evaluation is really impressive and a fantastic outcome. Combining this evidence together into formal-methods based reporting tools to build an all important assurance case completes the picture. Demonstrating that this can be done both at scale and in real-world systems is such a high impact indicator, and is something the field of formal methods needs substantially more of: well done.  
  
4. What do you see as the strongest aspect of this manuscript?: The paper surveys a lot of potential tools; the authors have selected, identified, and constructed a complete workflow at a full life-cycle of high assurance product evaluation. This produces a coherent, focussed workflow, and all of the approaches are fit for purpose for their proposed context. The paper embeds evidence into the formal assurance case throughout, generating, documenting, and structuring the evidence for the assurance case. This results in a nice mixture of formal verification tools and techniques in combination with formally verified technologies and components, e.g., seL4 to achieve the required assurance levels.  
  
5. What do you see as the weakest aspect of this manuscript?: Overall very good. However:  
- Difficulty in navigation of the paper: the paper presents a big toolset with lots of moving parts; lots of effort on the part of the reader to construct a mental model of what each of the individual things do and how they work together. I would like to see some 'high level', introductory summary and/or graphic of how all the parts relate.

Addressed above.  
  
- The paper would benefit from more transparency over a couple of areas: please give more discussion about real time requirements and operating systems, and please give more discussion about programming languages and environments. See comments below for more detail.

Addressed above.

ADDRESSED ABOVE

7. Please rate and comment on the timeliness and long term interest of this manuscript to S&P readers in the Detailed Comments section. Select all that apply.: Topic and content are of immediate and continuing interest to S&P readers  
  
Please rate the manuscript. Explain your choice in the Detailed Comments section.: Excellent