

From assessment to standardised benchmarking: Will it happen? What could we do about it?

Henrique Madeira¹ and Istvan Majzik²
¹ *CISUC, University of Coimbra, Portugal*
² *Budapest University of Technology, Hungary*
henrique@dei.uc.pt, majzik@mit.bme.hu

1. Introduction

Cost pressure, short time to market, and increased complexity are responsible for an evident increase of the failure rate of computing systems, while the cost of failures is growing rapidly, as a result of an unprecedented degree of dependence of our society on computing systems. The combination of these factors has created a dependability and security gap that is often perceived by users as a lack of trustworthiness in computer applications, and that is in fact undermining the network and service infrastructures that constitute the very core of the knowledge-based society.

Having effective and accurate methods and tools to assess dependability and security is essential to understand current risks of network and service infrastructures and contribute to improving the current situation. However, although considerable efforts have been made, assessing dependability and security is still a difficult problem. The quality of measurements, the assessment of dependability in component based, dynamic and adaptive systems and networks, and the integration with the development process are among the evident challenges. The problem is even harder when it comes to the assessment of dependability in a standard and comparable way, and all major classes of threats, namely accidental faults (component failures, software bugs, human mistakes, interaction mistakes) as well as malicious attacks are considered.

Thus it is time to ask the questions: **can we really expect to grow from dependability and security assessment to standardized benchmarking? Will it happen? What could we do about it?**

The panel is meant to foster lively debate looking at the main reasons why dependability and security assessment is still so difficult to attain in practice, even for relatively simple systems. Why dependability benchmarking still seems a distant promise? Is there stagnation in this research area? Or are there exciting recent developments? What future research directions seem promising? Is dependability/security benchmarking a consensual goal, or the uncertainties and dangers related to potential misuse of benchmarks recommend a cautious attitude? Who needs dependability/security benchmarks? Is there a clear business model to promote assessment and benchmarking activities?

2. Discussion focus

In addition to the general questions mentioned before, we have identified a set of dimensions of the problem that we use to motivate discussion:

- **Needs, drivers, and challenges:** What are the current needs and challenges in dependability and security assessment? What are the drivers for investments in improved assessment solutions? Among the challenges (e.g., unsolved problems created by the increasing demands from

users/regulators as well as by the continuing technological progress), which ones could be considered as *high priority challenges* that should be addressed as soon as possible? Which challenges need *long-term* study and research?

- **Recent developments:** What are the *recent developments* concerning the definition of (product-specific) metrics, assessment techniques, adoption of benchmarks that can be considered as leading solutions in this field? Are there any factors that hinder the use of the recent research results in the practice? Is research on dependability and security assessment really thriving (or is there a stagnation concerning new solutions)?
- **Future research directions:** What are the *promising research directions* that need attention and must be supported? How can the research community contribute towards an *agreement* that is represented by benchmarks (accepted by vendors and users)? How can it help to decide whether a potential benchmark is premature and possibly dangerous or not?
- **Training and education:** What are the *gaps* in engineering education regarding the measurement, assessment and benchmarking of dependability and security? What kind of *related training* is most needed?

The invited panelists are asked to focus their presentations (position statements) on one of the following areas of their expertise to solicit inputs from the audience and encourage debate (“brainstorming”):

- **Scientific foundations:** Advances that underlie the assessment and benchmarking technologies, including both *theoretically sound solutions* of relevant problems focusing on selected aspects, and the integration of partial solutions into a *holistic*

approach to address the complexity of systems and the (human, socio-economic) context in which systems operate.

- **Measurements:** Quantitative evaluation based on measurements and insights from *metrology* (the science of measurement) that are useful and exploitable to provide sound benchmarking principles and to increase our confidence in measurement results.
- **Assessment of dependability and security:** Work addressing important challenges and contributing towards (quantitative) *methods and tools* suitable for the current and next generation of systems.
- **Benchmarking and standardization:** Activities that can be seen as the evolution of current assessment techniques into more *standardized approaches* that offer generic, repeatable and widely accepted methods for characterizing the system behavior in the presence of faults, and allow the comparison of alternative solutions. This should be seen together with possible actions *fostering the standardization* related to assessment and benchmarking of dependability and security.

3. Organization of the Panel

The invited panelists cover the areas of expertise mentioned above and include experts from academia and industry, in particular experts from large international projects related to assessment and benchmarking of dependability and security and from specific groups such as the IFIP 10.4 Special Interest Group on Dependability Benchmarking.

Acknowledgment

The organization of the panel is supported by the AMBER - "Assessing, Measuring, and Benchmarking Resilience" Coordination Action project 216295 funded by the European Commission in FP7. The presentations of the panelists will be available at the project web portal:

<http://www.amber-project.eu/>