Report from the Artifact Evaluation Chair

21st Middleware Conference (Middleware 2020)

Tobias Distler

Friedrich-Alexander University Erlangen-Nürnberg (FAU)

Motivation

In an effort to encourage researchers to not only publish papers but also the associated prototype implementations, this year's Middleware conference offered authors of accepted papers the possibility to participate in an artifact evaluation process. The submission of artifacts was voluntary and the evaluation started after the shepherds approved the camera-ready versions. Consequently, the outcome of the artifact evaluation had no influence on paper acceptance decisions. Nevertheless, the authors of almost half of the papers accepted at Middleware 2020 opted to take part in the artifact evaluation.

Badges

Artifacts submitted for evaluation had the chance to receive one or both of the following badges recommended by ACM:

- **Artifacts Available (AA)**: This badge was awarded to artifacts that play a significant role in their respective papers and, more importantly, have been made publicly and permanently accessible.

- **Artifacts Evaluated – Functional (AE)**: A paper earned this badge if its artifact (1) contains sufficient documentation, (2) has been used to produce the presented measurement results, (3) includes (to the extent possible) all components relevant to the paper, and (4) comprises scripts/software that can be successfully executed.

Note that neither of these badges requires a reproduction or replication of the published results, and thus result validation was not part of the artifact evaluation process.

Evaluation Process and Outcome

At submission time, authors were given the opportunity to select the specific badge(s) they aim for. We received 14 applications for the AA badge, all of which were approved. 11 of those artifacts also underwent the more rigorous and time-consuming AE badge evaluation process, but only 8 passed it.

In total, the AE badge evaluation took four weeks during which every artifact was examined by one member of the evaluation committee. Specifically, reviewers were asked to check artifact completeness against the paper as well as to try to install and execute the implementation based on the instructions in the publicly available documentation. Typically, reviewers for this purpose used their own local systems, however one artifact required a specific hardware setup that was only available at the authors' institution. In this case the reviewer was granted remote access to perform the evaluation.

As an essential part of the evaluation process, reviewers were advised to repeatedly communicate with authors to give feedback on their findings and request assistance if they get stuck. This interaction was handled in a single-blind manner, that is without revealing the identity of reviewers to authors. The possibility to ask questions turned out to be crucial for the entire evaluation process. Without this tool (i.e., based on the initially submitted version alone), none of the papers would have received an AE badge. For three artifacts, the AE badge evaluation nevertheless was unsuccessful, either because there was not enough time to resolve all identified problems or due to the authors not responding to questions.

Conclusion

The continuous interaction and cooperation with reviewers enabled authors to significantly improve the quality and usability of their artifacts, as confirmed by both sides: reviewers and authors. Examples of issues fixed this way include incomplete execution instructions, missing dependencies, hard-wired host names and file-system paths, as well as programming errors. Hopefully, the updates will prove beneficial to other researchers interested in using the artifacts.

Thank you to the evaluation committee members for reviewing the artifacts, and to the authors for participating in the process and making their code publicly available!

Artifact Evaluation Committee

- Tobias Distler (*FAU Erlangen-Nürnberg*, chair)
- Michael Eischer (*FAU Erlangen-Nürnberg*)
- David Goltzsche (*TU Braunschweig*)
- Laura Lawniczak (*FAU Erlangen-Nürnberg*)
- Mohammad Mahhouk (*TU Braunschweig*)
- Manuel Nieke (*TU Braunschweig*)
- Nico Weichbrodt (*TU Braunschweig*)